Investing in AI Talent: Labor-Market Returns to AI Skilling College Infrastructure in India*

Yutong Chen† Sheetal Sekhri‡

This Version: March 2026

We study the labor-market effects of establishing AI centers at India's leading engineering institutions. Exploiting staggered adoption and recent difference-in-differences estimators, we find that AI center establishment increases full-time placement rates by about 4 percentage points and raises median earnings by 8-11 percent, with effects that strengthen over time. A simple conceptual framework highlights two margins, crossing a placement barrier and realizing wage gains conditional on entry, and the observed institution-level heterogeneity aligns closely with these mechanisms. AI centers also induce sharp expansions in computational infrastructure, faculty capacity, and research activity. The results show how targeted educational investments can shape labor-market outcomes in the AI era.

Keywords: Artificial Intelligence, Skills, Higher Education, Labor Markets
JEL Codes: I25, J24, O33, O15

*We are grateful to Milind Sohoni for discussions on the National Institutional Ranking Framework (NIRF), to Nivetha Thangaraj and Sparsh Badjate for their assistance in collecting data from the NIRF website, and to Jean Paul Rajesh for his assistance in collecting data from the Web of Science. We are also grateful to Sandip Sukhtankar for their insightful comments.

†Department of Economics, University of Texas at Arlington. Email: yutong.chen@uta.edu

‡**Corresponding Author**: Department of Economics, University of Virginia. Email: ssekhri@virginia.edu

1 Introduction

The rapid diffusion of artificial intelligence (AI) is part of a broader pattern in which frontier technologies reshape occupational structures, task content, and the returns to technical skills (Acemoglu et al., 2022; Hui et al., 2024). As firms adopt AI-intensive production technologies, the supply of workers able to complement these tools becomes a central determinant of labor-market performance. More broadly, however, new technological opportunities do not automatically translate into broad labor-market gains. Despite growing policy attention to AI-related education, there is limited causal evidence on whether higher-education institutions can meaningfully build AI-complementary human capital at scale—particularly in low- and middle-income economies, where mobility frictions and human-capital bottlenecks remain salient (Bryan et al., 2025) This paper studies whether technology-specific investments in tertiary education can help translate the global AI transition into early-career labor-market gains.

A motivating feature of the global landscape is the gap between countries with strong AI-related institutional capacity and those with rapidly expanding STEM supply but limited AI-specific infrastructure. Despite producing one of the world's largest cohorts of engineering graduates (around 1.5 million per year), India is characterized in the *State of Global AI Compute 2025* report as "compute-vulnerable," with limited installed compute capacity and substantial reliance on foreign hardware relative to the United States, China, and several upper-middle-income comparators. Complementary evidence from Maslej et al. (2025) points to similar gaps in the broader AI ecosystem: in 2023, India's granted AI patents were 2.6% of the U.S. total and effectively zero relative to China (rounded to 0.0%), and in 2024 private AI investment in India was 1.1% of the U.S. level and 12.5% of China's. These differences underscore a broader policy question: can investments in AI-complementary infrastructure meaningfully shape early-career labor-market outcomes?

We address this question by examining the establishment of dedicated AI centers at India's leading engineering institutions between 2017 and 2024. India provides a revealing setting for this analysis. It is home to one of the world's largest emerging engineering labor markets, and evidence points to rising employer demand for machine-learning and data-intensive skills. At the same time, industry and policy analyses suggest that the supply of such skills is not keeping pace, highlighting the importance of targeted investments in tertiary education and research capacity (Mani et al., 2020; World Bank, 2025). Recent micro-evidence further shows that AI-related demand has already begun to reshape wage offers and hiring patterns in Indian service-sector firms (Copestake et al., 2023).

The AI centers we study are institution-level investments that expand computational

laboratories, specialized faculty, and externally funded research in machine learning and data science, thereby increasing students' exposure to frontier tools and project-based training.[1] Their staggered rollout across institutions provides an opportunity to estimate the causal effects of concentrated expansions in AI-oriented educational capacity. Using institution-by-year administrative data from the National Institutional Ranking Framework (NIRF) and the Web of Science, together with recent difference-in-differences estimators designed for staggered adoption, we find that AI center establishment leads to persistent labor-market gains: undergraduate placement rates increase by approximately 4 percentage points, and median early-career earnings rise by 8-11 percent relative to non-adopting institutions. These effects strengthen over time, consistent with the gradual accumulation of instructional capacity and research-linked opportunities.

Beyond average effects, we document substantial institution-level heterogeneity that aligns with the conceptual framework in Section 5.2. On the extensive margin, placement gains are concentrated among institutions serving academically stronger cohorts, consistent with a *placement barrier* in which AI centers provide specialized infrastructure—including advanced laboratories, computational resources, and deeper faculty expertise—that complements baseline preparation and facilitates entry into AI-intensive jobs. On the intensive margin, conditional earnings gains vary systematically with cohorts' counterfactual opportunities and market access. Institutions serving more disadvantaged cohorts exhibit larger proportional earnings gains among placed graduates, consistent with a *wage-returns barrier* in which weaker outside options imply steeper proportional gains upon entry into AI-intensive roles. Separately, institutions serving more geographically mobile cohorts experience larger conditional earnings increases, consistent with market-access amplification: greater mobility is associated with better access to high-return AI labor markets and higher-quality matches once graduates are placed. We also find that treated institutions sharply expand computing laboratories, faculty capacity, and (potentially) industry-facing research activity, consistent with the institutional inputs emphasized by the framework. Taken together, these patterns provide institution-level evidence in line with canonical models of complementarity, sorting, and Roy-type heterogeneous returns.

[1]For example, leading engineering institutions in India have established interdisciplinary AI centers that combine advanced computational infrastructure, dedicated faculty, and externally funded research programs in machine learning and data science. These include the Robert Bosch Centre for Data Science and Artificial Intelligence at Indian Institute of Technology (IIT) Madras, which supports research and training in deep learning, natural language processing, and related areas, as well as AI research hubs at institutions such as International Institute of Information Technology (IIIT) Hyderabad and IIT Bombay that host specialized laboratories, faculty, and industry-linked research projects. Such centers are designed to integrate students into hands-on research and project-based learning using frontier AI tools (Ravindran et al., 2022).

Two alternative explanations are that our estimates reflect selective student redistribution or coincident local labor-market booms rather than the causal effect of AI center adoption. We find little evidence supporting either account. First, we do not observe statistically significant treatment effects on the share of within-state students, suggesting that cohort composition remains stable after AI center establishment. This pattern is consistent with India's centralized engineering admissions system, in which institution and major assignments are largely fixed at entry and across-institution transfers are extremely rare.[2] Second, our results are robust to restricting the analysis to a more selective subsample of top-ranked institutions and to specifications that include state fixed effects and flexible controls for district-level labor-market conditions, implemented by interacting baseline district employment, labor-force participation, unemployment rates, and pre-period changes in IT-related employment with year dummies. These controls absorb time-invariant regional differences and allow districts with different initial labor-market conditions to follow distinct time paths. Finally, recent evidence from online vacancy data indicates that the post-2016 surge in AI-specific labor demand in India did not translate into detectable district-level changes in total hiring or wage growth (Copestake et al., 2023), mitigating concerns that localized labor-market booms coinciding with AI center establishment drive our estimates. Taken together, these patterns strengthen confidence that the observed labor-market gains reflect the causal impact of AI center establishment rather than shifts in student composition or coincident local demand shocks.

Substantively, the paper advances knowledge in three areas. First, we contribute to a growing literature on artificial intelligence and labor markets. Prior work primarily uses task- or occupation-level measures of "AI exposure" to examine how advances in machine learning reshape employment and wages (Acemoglu and Restrepo, 2018; Acemoglu et al., 2022; Brynjolfsson et al., 2025a,b; Demirci et al., 2025; Webb, 2020). This research emphasizes that the effects of AI are highly skill-biased and depend critically on complementary investments in human capital and organizational practices. A parallel literature extends these exposure measures to low- and middle-income countries, showing how occupational structures shape the incidence of AI across development contexts (Demombynes et al., 2025). Recent work also documents patterns of firm-level AI adoption and shows that realized productivity and employment impacts are heterogeneous and often modest in the short run across four developed countries (Yotzov et al., 2026). This literature provides important insights into the demand side of AI adoption but leaves open a central question: how does the supply of AI-complementary human capital adjust to frontier technological change? We address this gap by shifting the focus

[2]See Bagde et al. (2016) for evidence on limited post-admission mobility in Indian engineering colleges.

from occupation-level exposure to the educational supply side, providing the first causal evidence on how targeted investments in AI-complementary educational capital affect early-career labor-market outcomes in a large emerging economy.

Second, we contribute to research on human-capital production and the distributional implications of skill-biased technological change. Prior work highlights complementarities between technology adoption and baseline skills (Krueger, 1993; Acemoglu and Autor, 2011) and the role of selection into high-return sectors (Heckman and Honoré, 1990), but evidence on how technology-specific educational investments shape these margins is scarce. We show that the returns to AI-oriented institutional investments vary systematically across institutions: those with stronger baseline preparedness realize larger placement gains, while conditional wage gains are larger for cohorts with weaker outside-option strength and stronger market access. These institution-level patterns extend canonical theories of complementarity and Roy-type selection to a frontier technological context and reveal how emerging technologies interact with existing educational inequality.

Third, we contribute to research on capability formation and technology adoption in developing economies. A broad literature emphasizes that participation in frontier sectors requires complementary investments in human capital and innovation infrastructure (Acemoglu et al., 2005; Cohen and Levinthal, 1989; Hausmann et al., 2007; Nelson, 1993). Most existing empirical work examines either broad expansions of STEM education or firm-side adoption of new technologies (Bloom et al., 2016; Goldin and Katz, 2018), and evidence from low- and middle-income countries remains largely descriptive. We provide the first causal evidence on a technology-specific investment in higher education—AI centers—that expands computational infrastructure, deepens faculty specialization, and increases sponsored research activity. These institution-level capacity improvements are followed by persistent gains in placement and earnings, indicating that targeted educational investments can help build the absorptive capacity needed for emerging economies to enter AI-intensive global value chains.

Overall, our findings show that the labor-market effects of frontier technologies are mediated by tertiary-education capacity. Technology-specific investments in higher education can expand access to AI-intensive jobs, but the gains are neither automatic nor evenly distributed: they depend on whether institutions can build the complementary skills required for entry and whether graduates can reach the markets where those skills command the highest returns. This framework helps explain why frontier technologies can simultaneously broaden opportunity and reinforce existing inequality. More broadly, the results highlight the central role of higher-education policy in shaping who benefits from technological change and underscore that building frontier-ready human capital re-

mains a first-order development challenge.

The remainder of the paper is organized as follows. Section 2 provides background on India's national AI strategy. Section 3 describes the data. Section 4 outlines the empirical design, including the staggered adoption of AI centers, the construction of institution-level panels, and the implementation of recent difference-in-differences estimators. Section 5.1 presents the results on placement and earnings; these patterns motivate a simple conceptual framework, which we develop in Section 5.2. The framework formalizes two barriers—placement and wage returns, and clarifies how institutional characteristics shape heterogeneous effects. We test the model's predictions in Section 5.3 and examine potential mechanisms in Section 5.4. Section 6 reports robustness checks. Section 7 discusses the study's limitations, and Section 8 concludes.

2 Background

India has rapidly expanded its engineering and computer-science education system over the past two decades, producing one of the largest cohorts of tertiary STEM graduates worldwide. However, AI-specific educational infrastructure, including high-performance computing resources, applied AI laboratories, specialized faculty, and industry-linked research programs, remains limited relative to the size of the student population and the growing demand from technology-intensive sectors.

Beginning in 2018, a combination of central government initiatives, state-level programs, and philanthropic efforts encouraged the creation of dedicated AI centers at selected engineering institutions. The 2018 "*AIforAll*" strategy identified AI-related capacity building as a national priority and supported the establishment of Centers of Excellence (CoEs) in emerging technology fields by providing funding. These CoEs complemented earlier state efforts and parallel initiatives by industry-funded foundations and higher-education philanthropies to set up new AI centers. As a result of different funding models, AI centers were introduced at different institutions in different years between 2017 and 2024, generating staggered adoption across institutions with similar baseline characteristics.

AI centers typically include: (i) computational laboratories with high-performance computing resources (see Figure 1a), (ii) faculty positions and training programs in AI and data science (see Figure 1b), (iii) industry-facing research projects, and (iv) curricular revisions expanding AI-oriented courses and experiential learning opportunities. These components mirror the forms of "AI-complementary educational capital" emphasized in our conceptual framework in Section 5.2.

Temporal Distribution Figure 2 presents the distribution of establishment years across institutions, with the blue bars showing these for the 80 institutions with confirmed years.[3] There is a clear surge in center establishment after 2018, consistent with institutional responses to the National Government's *AIforall* Strategy. The number of new centers peaks in 2022, when 18 institutions established an AI center.

Geographical Distribution Figure 3 displays the geographic distribution of the 113 institutions included in the main analysis. The institutions are spread across 21 states and union territories, with some concentrations in southern states such as Kerala and Tamil Nadu, as well as around the Delhi-NCR region. Institutions that established an AI center between 2017 and 2024 are indicated by colored markers, while those without such centers are shown in gray. The distribution reveals that the establishment of AI centers spans diverse regions rather than being concentrated in a few states, reflecting the nationwide diffusion of institutional participation.

We leverage this temporal and geographic variation in the establishment of AI centers in our identification strategy, as discussed in detail in Section 4. Additional policy background is provided in Appendix B, and Appendix Table B8 lists the institutions along with the corresponding years in which their AI centers were established.

3 Data

We combine multiple administrative and hand-collected sources to construct the first micro-linked dataset on AI center adoption and graduate outcomes in India. Our primary data come from the National Institutional Ranking Framework (NIRF), which provides detailed performance statistics for academic institutions. We hand-collected information on the establishment and timing of AI centers and merged it with the NIRF data.

3.1 AI Centers in Indian Engineering Institutions

To identify whether and when an institution established an AI center, we manually collected information from multiple sources. For each institution, we reviewed official websites, social media announcements (e.g., Facebook and LinkedIn), and coverage in major

[3] This data was collated by us and we could confirm dates for 80 among 87 AI centers. See Section 3 and Appendix C for details.

news outlets such as *The Hindu* and *The Times of India*.[4] We cross-validated the dates across these sources, which provide consistent documentation of both the existence and timing of AI center establishment. Among the 268 institutions in our sample, we identify 87 that had established an AI center by 2024. For seven institutions, the establishment year could not be verified, and we exclude these institutions from the analysis.[5]

3.2 NIRF and Web of Science Institution-by-Year Panel Data

Our analysis uses institution-by-year administrative data from NIRF, covering engineering institutions in India from 2016 to 2024.[6] The data contain rich information on student outcomes, faculty and research inputs, and financial resources. We construct an institution-level panel using annual information on student placement rates, median earnings among placed graduates, the percentage of students pursuing higher studies, capital and operating expenditures, and student demographics. We focus primarily on outcomes for four-year undergraduate programs, which represent the modal engineering degree and account for the largest share of graduates entering technical labor markets.

The NIRF data are compiled through annual submissions by institutions and are collected to construct rankings as a metric of quality. While institutions certify the accuracy of their reported information, NIRF performs internal consistency checks and validates submissions against regulatory data. Research-related indicators are independently verified using third-party bibliometric databases such as Scopus and the Web of Science. In addition, these submissions are subject to public audit.[7] Appendix C provides details on how these data are assembled.

We further supplement the NIRF panel with bibliometric data from the Web of Science to capture institution-level AI research output. Specifically, we construct institution-by-year measures of AI-related publications affiliated with each engineering institution in

[4]Funding sources are not consistently disclosed for AI centers: only a small subset of institutions explicitly name sponsors or donors on public webpages (e.g., IIT BHU reports support from Coforge). As a result, we cannot systematically classify centers by funding origin for the sample.

[5]We assess whether these institutions differ systematically from those for which we were able to verify the year of establishment in Appendix Table B9. We find that the two groups are not jointly statistically significantly different across baseline characteristics at conventional levels, including overall enrollment, student composition, and expenditure measures.

[6]Note that the 2017 rankings are based on statistics from the 2015-16 academic year. Accordingly, our dataset incorporates the rankings published from 2017 to 2025. We provide additional details on data construction in Appendix C.

[7]Under NIRF guidelines, institutions are required to make their most recent submission publicly accessible, enabling external validation and enhancing transparency in the reporting process.

our sample and aggregate these records annually.[8] These measures complement NIRF's research indicators, such as sponsored projects and related funding, by capturing realized scholarly output from an externally compiled source. They therefore provide a useful additional measure of institutional research capacity and academic visibility.

3.3 Main Outcomes

Our primary outcomes of interest are (i) the placement rate, defined as the ratio of students placed to students graduating; (ii) the higher studies rate, defined as the ratio of students pursuing higher studies to students graduating; and (iii) the median annual salary of placed graduates. To ensure comparability over time, we normalize all expenditure variables (e.g., capital expenditure on library resources) and the median annual salary using the Consumer Price Index (CPI), with 2016 as the base year. All monetary values are therefore expressed in real 2016 INR.

3.4 Sample for Main Analysis

For our main analysis, we restrict the sample to institutions with an average NIRF engineering rank of 100 or better for two reasons.[9] First, top-ranked institutions are more comparable in scale, resources, and academic reputation, and thus are better suited for constructing counterfactuals for institutions that established an AI center.[10] Second, consistent data across all years are only available for institutions ranked 100 or better.[11] Importantly, an institution's 2017 rank is not correlated with subsequent adoption of an AI center, making this restriction innocuous.[12]

Our main analysis sample includes 113 institutions, of which 34 established an AI center between 2017 and 2024. Because some institutions share the same rank, the total exceeds 100 even though we restrict the sample to institutions ranked 100 or better.[13] We also use the full sample of all 268 institutions to assess the robustness of our results in

[8]The keywords we used include artificial intelligence, machine learning, deep learning, neural networks, large language model, LLM, and generative AI. We restrict attention to articles and conference proceedings and exclude retracted publications.

[9]Appendix C describes how NIRF ranks are computed.

[10]Indian engineering sector saw a steep surge in private engineering institutions in the last 2 decades but the quality of these new institutions is not at par with the top-ranked institutions.

[11]See Appendix C for details.

[12]We regress an indicator for whether an institution established an AI center between 2017 and 2024 on its 2017 NIRF engineering rank. Note that the 2017 ranking is computed using data from the 2015-2016 academic year. The estimated coefficient is -0.002 ($p = 0.25$).

[13]We exclude one institution that had already established an AI center by 2016, as it is always treated and lacks a pre-treatment period for comparison.

Section 6. As in the full sample (discussed in the Background section), there is variation in the timing of the establishment of AI centers among our main analysis sample. The yellow bars in Figure 2 show the distribution of the establishment dates of the 34 AI centers in this main analysis sample. We observe a surge in establishment activity after 2018, analogous to the full sample.

3.5 Descriptive Statistics

Table 1 reports descriptive statistics for the main outcome variables between 2016 and 2018.[14] For four-year undergraduate programs, the average placement rate is 70.6%, and 11.7% of graduates pursue higher studies. The average annual median salary of employed undergraduates is 585,362 INR (in real 2016 terms). This is equivalent to 8,722 USD in 2016.[15] For two-year master's programs, the average placement rate is lower at 48.4%, and the average annual median salary is 556,000 INR (8,284 USD in 2016). In terms of expenditures, institutions spend, on average, 114 million INR (1.7 million USD in 2016) on new equipment and software for laboratories, 582 million INR (8.7 million USD in 2016) on salaries for teaching and non-teaching staff, and 408 million INR (6.1 million USD in 2016) on maintenance. Finally, in 2018, institutions undertook an average of 93 sponsored research projects receiving 226 million INR (3.4 million USD in 2016) in total funding, and 141 consultancy projects earning 55 million INR (0.8 million USD in 2016).

We also report descriptive statistics comparing institutions that established an AI center between 2017 and 2024 with those that did not in Appendix Table C11. These institutions differ along several dimensions, including student composition and multiple categories of capital and operating expenditures. Our empirical strategy addresses these differences and constructs plausible counterfactuals to identify causal effects.

4 Econometric Framework and Identification

Our empirical goal is to estimate the causal effect of establishing an AI center on student labor-market outcomes. We exploit staggered variation in the year of establishment across institutions between 2017 and 2024. The treatment timing corresponds to the first year an institution reports an operational AI center. Because treatment is adopted at different times, we implement three recent Difference-in-Differences (DiD) estimators designed for staggered-adoption settings with potentially heterogeneous treatment effects across

[14] Research project variables are shown only for 2018, the first year in which they become available.
[15] The average exchange rate in 2016 was 1 INR = 0.0149 USD.

cohorts. We interpret all estimates as institution-level Average Treatment Effects on the Treated (ATT). Our preferred specification is Synthetic Difference-in-Differences (SDID), while we use the imputation estimator of Borusyak, Jaravel and Spiess (2024) (BJS) as a complementary robustness check.

4.1 Synthetic Difference-in-Differences (SDID)

We first implement the Synthetic Difference-in-Differences (SDID) estimator of Arkhangelsky et al. (2021). In our setting, AI-center adoption is unlikely to be random: as discussed in Section 3.5, treated institutions differ from never-treated institutions along several dimensions, raising a concern that conventional DID may not provide a credible counterfactual. SDID addresses this concern by combining unit-level synthetic control weighting ($\hat{\omega}_i$) with time weighting ($\hat{\lambda}_t$) to construct a comparison group that flexibly absorbs common shocks and closely matches the pre-treatment trajectories of treated institutions. The estimator $\hat{\tau}^{\text{sdid}}$ is derived from a regularized least-squares problem:

$$(\hat{\tau}^{\text{sdid}}, \hat{\omega}_i, \hat{\lambda}_t) = \arg\min_{\tau,\mu,\alpha,\gamma} \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{\omega}_i \hat{\lambda}_t (Y_{i,t} - \mu - \alpha_i - \gamma_t - \tau D_{i,t})^2. \tag{1}$$

Here, $Y_{i,t}$ denotes the outcome of institution i in year t, μ is a constant intercept, and α_i and γ_t represent institution and year fixed effects, respectively. $D_{i,t}$ denotes the treatment indicator for institution i in year t. Standard errors are computed using a bootstrap procedure with 999 replications.[16] While the exposition above presents the canonical formulation, our empirical setting involves staggered treatment adoption, which requires a slightly modified implementation; we provide a detailed discussion in Appendix D.1.

Data Requirement For the SDID estimator, we construct a balanced panel of institutions with complete observations over the entire 2016–2024 period. This requirement ensures the optimal weights are consistently identified across all time periods, a crucial design feature of the SDID procedure.

4.2 Borusyak, Jaravel, and Spiess (BJS) Imputation

To complement the SDID results and relax the requirement of a balanced panel, we also apply the imputation estimator of Borusyak, Jaravel and Spiess (2024) (BJS), which exploits the full set of available institution-year observations. This approach constructs

[16] We additionally report placebo-based standard errors from 999 replications as a robustness check. See Appendix D.2 for details.

counterfactual outcomes for treated units by first estimating the underlying two-way fixed effects ($\hat{\alpha}_i$ and $\hat{\gamma}_t$) using only non-treated institution-year observations (that is, observations with $D_{i,t} = 0$, including both never-treated institutions and pre-treatment years of treated institutions). The ATT is then calculated as the average difference between the observed outcome and the imputed counterfactual, $\hat{Y}_{i,t}(0) = \hat{\alpha}_i + \hat{\gamma}_t$. This method is efficient and relies on the parallel-trends assumption for untreated potential outcomes implied by the factor model. We report standard errors clustered at the institution level for the BJS estimates.

4.3 Identification and Robustness

Identification across all three methods ultimately relies on the assumption that, conditional on the method-specific controls (weights, fixed effects, or factor structure), treated and comparison institutions would have exhibited similar trends in the absence of AI center establishment. The estimand is the institution-level average treatment effect on the treated (ATT).

Evidence on Parallel Trends The parallel-trend assumption receives strong support in the data. Figure 5 and Appendix Figure E1 report event-study estimates from the SDID[17] and BJS methods for undergraduate placement and earnings and operating spending on salaries. These figures show mostly flat and statistically indistinguishable pre-treatment trends across outcomes. Together, these patterns strengthen the case that treated and control institutions were on comparable trajectories prior to AI-center establishment.

Adoption Timing A key concern is that AI-center adoption timing may be endogenous. As a complementary diagnostic, we test whether the *timing* of adoption is systematically related to observable pre-treatment characteristics. Specifically, we regress institutions' first adoption year on baseline (2016) measures of undergraduate scale (total four-year undergraduate enrollment) and student composition (female share, within-state share, and the share of students from socially disadvantaged groups), capital expenditures (library; laboratory equipment and software; engineering workshops; other assets), and operating expenditures (salaries for teaching and non-teaching staff; maintenance; seminars). These baseline covariates are not jointly predictive of adoption year ($p = 0.46$).

[17] Under staggered adoption, SDID solves a separate optimization problem for each adoption cohort, yielding cohort-specific unit and time weights and thus cohort-specific synthetic counterfactuals. In our setting, SDID achieves close pre-treatment fit by placing weight on control institutions with similar pre-treatment trajectories. See Appendix Tables D12-D13 for the distribution of synthetic weights used in the SDID estimation for four-year undergraduate program outcomes.

Restricting the sample to the top-100 institutions yields similar results ($p = 0.75$). Together, these tests suggest limited evidence that adoption timing is related to observed baseline characteristics.[18]

We further assess potential selection into adoption timing using a discrete-time adoption hazard model and an institution-level predicted adoption-risk score. The results align with the timing tests above and provide no evidence of differential pre-treatment trends along the predicted-risk dimension (Appendix Table E1).[19]

Scope and Remaining Threats We note a scope restriction due to the data's aggregation: we cannot estimate micro-level mechanisms or individual-level responses. All heterogeneity analyses rely exclusively on pre-treatment institution characteristics, which avoids mechanical biases but limits individual-level interpretation. We address the possibility of cross-institution spillovers in Section 6.

5 Empirical Results

In this section, we first examine the effects of AI center establishment on labor-market outcomes for graduates of four-year undergraduate programs and two-year master's programs. We then explore potential mechanisms that may contribute to the observed labor-market effects.

5.1 Average Effects of AI Center Establishment

Undergraduate Programs Establishing an AI center generates sizable improvements in labor-market outcomes for graduates of four-year undergraduate programs.[20] Across the SDID and BJS estimators (columns 3-4 in Table 2), the share of graduates obtaining full-time employment rises by roughly 4 percentage points ($p < 0.05$), corresponding to an

[18] Given the modest sample size, we interpret this test as descriptive evidence on selection along observables rather than a definitive test of exogeneity.

[19] We estimate a discrete-time adoption hazard using baseline (2016-2018) covariates with year and state fixed effects and construct an institution-level predicted adoption-risk score from the fitted linear predictor. In addition to the baseline covariates described above, we include each institution's NIRF rank. We then estimate institution-specific pre-treatment slopes by regressing each outcome on calendar year using only "clean" pre-treatment years (at least two years prior to adoption for treated institutions; 2016–2018 for never-treated institutions) and regress these slopes on the predicted risk score. Predicted adoption risk does not forecast pre-treatment slopes in four-year undergraduate placement rates ($p = 0.75$) or median salaries ($p = 0.55$).

[20] The number of observations varies across columns because the SDID estimator is implemented on a balanced panel, which restricts the sample to institutions with complete data over the analysis period.

increase of about 5-6 percent relative to the 2016 baseline mean.[21] Under the SDID (or BJS) estimator, this effect corresponds to an increase of approximately 0.29 (or 0.25) standard deviations in the placement rate, indicating a quantitatively meaningful improvement at the institution level.[22]

Earning outcomes exhibit similarly positive gains. Median earnings among employed graduates increase by 8-11 percent, with estimates consistently positive and statistically significant at conventional levels (columns 5-6). These earnings gains are economically meaningful, albeit modest. Because all salary measures are deflated to constant 2016 INR, the effects represent real increases in purchasing-power-adjusted earnings.[23] The SDID (or BJS) estimate in column 5 (or 6) implies an increase of approximately 51.6 (or 66.5) thousand INR relative to a 2016 baseline median salary of 605 (or 577) thousand INR, an improvement of roughly 8% (11%) in the first placement cycle. This corresponds to about 0.19 (or 0.24) standard deviations of the untreated outcome variable, placing the magnitude in the medium-effect range.[24] AI centers primarily operate on the extensive margin (placement), with modest but detectable effects on entry-level earnings.

These effects emerge gradually rather than as a one-time jump. Event-study estimates (Figure 5) show that placement gains appear in the first post-establishment year and continue to accumulate over subsequent cohorts. Earnings effects follow a similar trajectory, becoming more pronounced as AI centers mature and institutional investments are integrated into instruction and recruitment pipelines. While earnings effects are noisier than placement effects, the event-study estimates show little evidence of pre-trends and suggest gradual institutional adjustment rather than an immediate wage response. Finally, we find little evidence that AI-center establishment crowds out the pursuit of higher education: the SDID and BJS estimates for the share of graduates entering further study are small in magnitude and statistically indistinguishable from zero.

[21] As placement is a rate outcome, proportional changes provide a natural summary of improvements along this margin. Consistent with this interpretation, log specifications of the placement rate—reported in Appendix Table E2—show statistically significant relative increases in employment.

[22] The standard deviation of the placement rate among institutions without AI centers is 0.15.

[23] We focus on real earnings levels rather than log specifications. Median salary is an institution-level outcome with limited proportional dispersion; as a result, log transformations substantially compress variation and reduce statistical power. Level estimates therefore, provide a more transparent and economically meaningful measure of income differences at labor-market entry. Nonetheless, we report results using log specifications for both the placement rate and median salary in Appendix Table E2. The log placement estimates confirm that AI center establishment increases the share of graduates obtaining employment, with effects that are statistically significant across estimators and consistent with the level results reported in the main text. In contrast, log salary estimates are smaller in magnitude and imprecisely estimated, reflecting the limited proportional variation in institution-level median salaries. The estimated signs remain consistent with the level-based results.

[24] The standard deviation of the median salary among institutions without AI centers is 275.94.

Robustness Check We additionally implement the non-parametric estimator of Callaway and Sant'Anna (2021) (CS) for the undergraduate placement and salary outcomes. This approach directly estimates group-time average treatment effects ($\tau_{g,t}$) for each cohort g in each period t. In our setting, the CS estimator identifies these effects by comparing cohort g to a valid control group consisting of never-treated institutions and not-yet-treated observations from treated institutions. One advantage of this approach is that it avoids the negative-weighting problem that can arise in two-way fixed-effects estimators. Identification relies on the group-specific parallel trends assumption, which is weaker than the global parallel trends assumption underlying the standard two-way fixed-effects model. We then aggregate the $\tau_{g,t}$ estimates to obtain both the overall pooled ATT and the event-study profiles. Inference is based on robust asymptotic standard errors clustered at the institution level. Appendix Table E3 reports the corresponding estimates. The CS results are consistent with our main findings. Undergraduate placement increases by 4.1 percentage points and is statistically significant, closely matching the magnitudes obtained using SDID and BJS. The salary estimate is also positive, although it is imprecisely estimated.

Master's Programs In contrast, we detect no measurable short-run effects on labor-market outcomes for graduates of two-year master's programs (Table 3). Across all three estimators, the effects on higher studies, employment, and median earnings are small and statistically insignificant. A plausible interpretation is that four-year programs provide substantially greater exposure to the new AI-related infrastructure, coursework, and research activity. Master's students, by comparison, typically arrive with a shorter time to graduation and weaker integration into undergraduate-focused AI center initiatives, limiting their opportunities to convert these new inputs into placement and wage gains. This pattern is consistent with our framework: the longer the exposure to AI-complementary capital, the more likely students are to cross the placement barrier and benefit from subsequent higher earnings levels.

5.2 Conceptual Framework

The average effects provide a parsimonious view of the impact of AI center establishment, but they can mask substantial differences across institutions. To interpret these heterogeneous responses and the mechanisms underlying them, we develop a simple conceptual framework in which AI centers relax two barriers to success in AI-intensive careers. The framework is intentionally parsimonious and focuses on institution-level channels that

our data can support. Figure 4 displays the framework.

The first channel is a *placement barrier*. Students must attain a threshold level of AI-complementary skills to be competitive for AI-intensive roles. Investments associated with an AI center—including expanded laboratory capacity, enhanced computational resources, and deeper faculty expertise—increase the efficiency with which students convert baseline preparation into AI-complementary skills. In this sense, AI centers raise the productivity of the human-capital production function by providing inputs that are complementary to students' pre-existing preparation. Institutions serving cohorts with stronger baseline preparedness may therefore experience larger placement gains because they are better positioned to translate AI-center investments into employable skills. This mechanism is consistent with theoretical models of skill-biased technological change, in which new technologies complement pre-existing skills and raise the returns to more prepared workers (Acemoglu, 1998; Acemoglu and Autor, 2011; Krueger, 1993).

The second channel is a *wage-returns barrier*. Even conditional on entering AI-intensive occupations, students may realize different gains for two distinct reasons. First, proportional gains depend on counterfactual opportunities in the absence of an AI center. When outside options are strong, entry into AI-intensive jobs represents a smaller departure from the counterfactual career path, yielding more modest proportional wage gains. This is consistent with Roy-type occupational choice models, in which individuals with weaker counterfactual opportunities experience steeper gains upon entering high-return sectors (Heckman and Honoré, 1990; Roy, 1951). Second, conditional wage gains may be amplified by access to high-return labor markets and high-quality matches: cohorts with greater geographic mobility can more readily reach high-paying AI jobs once placed, so the returns to AI-complementary skills are larger when market access is stronger.

Together, these channels provide a structured interpretation of our heterogeneous treatment effects: baseline preparedness governs how strongly AI centers relax the placement barrier, while conditional wage gains vary with both outside options and market access. Appendix G presents a stylized institution-level model that formalizes these mechanisms and derives the corresponding heterogeneity predictions.

Pre-treatment Characteristics Our heterogeneity analysis uses pre-treatment institutional characteristics that proxy for baseline preparedness, outside options, and market access. These measures are supported by well-documented empirical relationships in India:

- **Social disadvantage (Scheduled Caste share)**: Higher shares of socially disadvantaged (Scheduled Caste) students are associated with lower average pre-college

preparation and more limited access to professional labor markets (Desai et al., 2010; Hnatkovska et al., 2012; Muralidharan and Prakash, 2017). We therefore use SC share as a proxy for *lower baseline preparedness* and *weaker outside options*.

- **Geographic mobility and market access (in-state share):** Geographic composition is informative about both selectivity and mobility frictions that shape access to high-return labor markets. Institutions drawing more students from outside the home state tend to be more selective and to attract academically stronger applicants (Bagde et al., 2016). Conversely, a higher in-state share suggests stronger local anchoring and mobility frictions, which may limit access to high-wage urban markets (Hnatkovska and Lahiri, 2013). We therefore use the in-state share as a proxy for *lower geographic mobility* and *weaker market access* (and, potentially, a less selective applicant pool).

These pre-treatment differences motivate the heterogeneity analysis that follows, where we examine how the effects of AI center establishment vary across institutions facing different degrees of the placement and wage-returns barriers.

5.3 Heterogeneous Effects

To implement the conceptual framework empirically, we stratify institutions along two pre-treatment dimensions: (i) cohort socioeconomic disadvantage and (ii) cohort geographic mobility. For each dimension, we split institutions above and below the sample median and estimate post-treatment effects using the SDID and BJS estimators. Table 4 reports the resulting estimates.

Placement We begin with the extensive margin. The framework predicts that placement gains should be larger where the *placement barrier* is easier to relax, namely at institutions with stronger baseline preparedness.

The split by socioeconomic disadvantage is consistent with this prediction. Columns 1-2 of Panel A of Table 4 show that institutions serving a lower share of socially disadvantaged students exhibit larger gains in full-time placement, while institutions serving a higher share show smaller and less precisely estimated effects. The "Treatment (Below)" estimates for placement are positive and statistically significant across both estimators, and the corresponding difference is negative and statistically significant under the BJS estimator. These patterns suggest that institutions serving less disadvantaged cohorts are better positioned to convert AI-complementary educational capital into successful entry

into AI-intensive jobs, consistent with the *placement barrier* being more readily relaxed where baseline preparedness is stronger.

By contrast, the split by geographic mobility yields weaker heterogeneity in placement. Columns 3-4 of Panel A show that institutions with lower within-state shares have positive and statistically significant placement effects, whereas the corresponding estimates for institutions with higher within-state shares are statistically insignificant. However, the difference in placement effects across the two groups is small and statistically indistinguishable from zero under both estimators. The results suggest that geographic mobility, which we interpret primarily as a proxy for market access, is less central for the extensive-margin placement response than to the returns realized after placement.

Wage We next turn to the intensive margin. The framework predicts that conditional wage gains should vary with both outside-option strength and market access.

The socioeconomic-disadvantage split reveals a distinct pattern on the intensive margin. Institutions serving a higher share of socially disadvantaged students show limited evidence of increased placement, yet placed graduates realize larger proportional earnings gains. In columns 1-2 of Panel B of Table 4, the "Treatment (Above)" estimates for median real salary are positive and statistically significant under the BJS estimator, and the corresponding difference estimates show a similar pattern. This is consistent with the *wage-returns barrier* and the comparative-advantage implication of Roy-type occupational choice models: cohorts facing weaker counterfactual opportunities can exhibit larger proportional wage gains conditional on entry, even when the probability of entry itself changes only modestly.

The geographic-mobility split points to a related but distinct intensive-margin pattern. Institutions serving more geographically diverse incoming cohorts—that is, those with lower within-state shares—exhibit larger proportional earnings gains among placed graduates, even though differences in placement rates are modest. The "Treatment (Below)" estimates for median earnings are positive in both specifications (columns 3-4 of Panel B) and statistically significant under the BJS estimator. The difference between "Treatment (Above)" (institutions with more in-state students) and "Treatment (Below)" is negative and statistically significant under BJS. In the context of our framework, this pattern is consistent with market-access amplification: conditional on placement, higher mobility improves access to high-return AI labor markets and higher-quality matches, increasing the returns to AI-complementary skills.

To summarize, the heterogeneity results exhibit a coherent two-stage pattern. On the extensive margin, placement gains are concentrated among institutions with stronger

baseline preparedness, consistent with complementarity between AI-center investments and pre-existing skills. On the intensive margin, conditional wage gains vary systematically with outside options and market access: historically disadvantaged cohorts can experience larger proportional gains relative to weaker counterfactual opportunities, while more geographically mobile cohorts realize larger earnings increases through improved access to high-return AI labor markets and higher-quality matches. Together, these findings support the interpretation that AI centers relax the placement and wage-returns barriers in systematically different ways across institutions.

Gender We additionally examine heterogeneity by gender. We find that institutions with larger baseline shares of women do not experience a statistically significant placement advantage (see the differences between "Treatment (Above)" and "Treatment (Below)" in columns 3-4 of Appendix Table E4). A striking result is that these institutions exhibit an earnings disadvantage. The differences in real median earnings are negative in columns 5-6, and the effect is statistically significant under the BJS estimator (column 6). This pattern is consistent with the large gender wage gap observed in STEM fields and with women facing wage penalties relative to men (Kahn and Ginther, 2017; Sahoo and Klasen, 2021). The expansion of AI-complementary education does not appear to mitigate this gap, at least in the short run.

5.4 Mechanisms: Institutional Capacity and Research Activity

We next examine how AI center establishment alters the internal capacity of institutions, focusing on changes in AI-complementary educational capital that operate through the two channels in Section 5.2. Across three domains—computational infrastructure, instructional staffing, and sponsored research—the introduction of an AI center generates targeted institutional expansions that map naturally onto the placement and wage-returns barriers.

Laboratory and Computational Infrastructure The strongest institutional adjustment occurs through expanded digital and computational infrastructure. Columns 1-4 of Table 5 show sharp increases in capital spending on laboratory equipment and software, nearly doubling relative to 2016 baseline levels (column 2). These investments enhance students' exposure to hands-on AI applications and high-performance computing environments that are central to AI-related instruction, while spending on engineering workshops, libraries, and other capital assets remains unchanged or even declines. This pattern suggests that improvements are concentrated in domains most closely aligned with

AI curricula and recruiter demand. Within the two-barrier framework, these targeted upgrades raise the productivity of AI-complementary educational capital and relax the placement barrier. Institutions with higher baseline preparedness are better able to translate these investments into employable skills. As a result, they experience larger placement gains; in our setting, baseline preparedness is proxied by serving more academically selective or more geographically diverse cohorts, consistent with the heterogeneity results.

Faculty and Technical Staffing Column 5 of Table 5 show sizable increases in spending on faculty and technical personnel. Event-study evidence in Appendix Figure E1 indicates a similarly gradual upward pattern in salary expenditures after AI center introduction. These investments deepen advanced instruction, ensure reliable access to updated computational environments, and support sustained project-based learning. Such expansions further relax the placement barrier for cohorts with stronger baseline preparedness, who are able to internalize new content more efficiently. At the same time, enhanced instructional and technical support broadens access to advanced skill-building opportunities for students at institutions where weaker outside options are more prevalent or market access is smaller, contributing to the conditional wage gains documented in Section 5.3.

Sponsored Research and External Engagement A third channel operates through externally funded research and industry engagement. Columns 1–3 of Table 6 show significant increases in sponsored research activity,[25] which can support faculty-led projects, create research-assistant opportunities, and raise institutional visibility in frontier AI domains. Sponsored projects may also deepen industry collaboration—through applied research partnerships, joint supervision, and informal recruiting pipelines—thereby strengthening links between curricula and employer needs.

We find complementary evidence of research expansion using bibliometric data on AI-related publications from the Web of Science. Table 7 reports positive treatment effects across estimators, with the strongest evidence in the BJS specification and in the full sample. On average, institutions publish about 13 more AI-related papers annually after establishing an AI center than comparable non-treated institutions (column 1, $p < 0.10$). These publication results complement the sponsored-project findings by capturing realized scholarly output rather than institution-reported research activity alone. Taken together, the evidence suggests that AI centers do not simply relabel pre-existing activ-

[25] By contrast, we do not observe a similar pattern for consultancy projects in columns 4-6.

ity; they are associated with a broader strengthening of institutions' AI research capacity, academic visibility, and external linkages.

Sponsored research reinforces both channels: on the placement margin, active research agendas signal program quality to recruiters and attract more mobile and high-ability students; on the wage-returns margin, experiential learning through sponsored projects improves job-match quality once students enter AI-intensive roles, helping explain the steeper proportional earnings gains observed for socially disadvantaged and more geographically mobile cohorts.

Discussion Overall, the institutional adjustments triggered by AI center establishment across digital infrastructure, staffing, and research engagement align closely with the two-stage conceptual framework. Expanded AI-complementary educational capital relaxes the placement barrier most strongly at institutions with higher baseline preparedness (i.e., a smaller share of socially challenged undergraduate students). At the same time, enhanced access to frontier learning opportunities relaxes the wage-returns barrier for cohorts facing weaker outside options (i.e., a higher share of socially challenged undergraduate students) and for cohorts with greater geographic mobility. Together, these adjustments provide a unified mechanism linking investments in AI-complementary educational capital to the differential placement and wage effects estimated above, strengthening the causal interpretation of our findings.

6 Robustness Checks

In this section, we assess the stability of our findings through a series of robustness checks. We begin with a permutation-based placebo test that evaluates whether staggered treatment timing could mechanically generate spurious effects. We then conduct checks that address alternative explanations related to sample composition, institutional trends, and geographic economic conditions.

6.1 Placebo Test

We conduct a permutation-based placebo test by randomly assigning placebo adoption years to institutions that never established an AI center. For each draw, we re-estimate treatment effects using the BJS methods 1,000 times for undergraduate placement rates and median salaries. In each iteration, placebo treatment years are drawn from the empirical distribution of actual adoption years to preserve the observed staggered timing

structure. Column 4 of Appendix Table F1 reports the empirical p-values, defined as the share of placebo effects that are at least as large in absolute value as the baseline estimate, which remain statistically significant at conventional levels. This provides additional support that the observed effects are not driven by spurious timing patterns or random fluctuations in untreated institutions.

6.2 Are Results Sensitive to Sample Restrictions?

We further extend the analysis to all institutions with available NIRF data. Because SDID requires a balanced panel—and only one institution ranked below 100 is observed in all nine years—we rely on the BJS estimator for this exercise. The full-sample results (Appendix F.2) remain consistent with our main findings in direction and statistical significance. This confirms that our conclusions are not driven by restricting the sample to institutions with complete balanced-panel data or to those with an average engineering rank of 100 or better.

6.3 Independent Trends Among Elite Institutions

One potential concern is that treated institutions may have experienced rising placement and earnings outcomes even in the absence of AI centers, particularly if they were already high-performing. To assess this possibility, we re-estimate all specifications using a sample restricted to institutions with an average NIRF engineering rank of 50 or better. This approach limits comparisons to a group of elite institutions while preserving staggered treatment timing.[26] As shown in Appendix F.3, the estimated effects remain similar in sign, magnitude, and statistical precision to the baseline results. Moreover, AI centers are not disproportionately concentrated among top-ranked institutions: as shown in Appendix Table B8, about 55% of adopting institutions have an average rank above 100, alleviating concerns that elite-specific trajectories drive our findings.[27]

6.4 Do Students from Control Institutions Transfer to Treated Ones?

A further concern is that our estimates may reflect selective transfers of students from non-treated to treated institutions following AI center establishment. Two pieces of evidence mitigate this concern.

[26] Adoption of AI centers is not correlated with baseline NIRF rank in 2016.
[27] In each year, institutions with both higher and lower average ranks adopt AI centers.

First, structural features of Indian engineering admissions substantially limit cross-institution movement. Nearly all seats in accredited four-year engineering programs are allocated through centralized entrance examinations, most prominently the Joint Entrance Examination (JEE) and state-level engineering entrance tests—followed by a single-round matching process that assigns students to institutions and programs based on rank and seat availability.[28] There is minimal scope for transfers after enrollment, and lateral admission into higher-ranked institutions is extremely rare. These institutional frictions make post-treatment sorting across colleges an unlikely explanation for our results.

Second, we assess whether the results could be driven by treatment-induced redistribution of students from non-treated to treated institutions. Specifically, we apply the SDID and BJS estimators to the share of within-state students. Appendix Table F8 reports the estimated ATT effects, which are small and statistically insignificant across all three methods. The findings suggest that the treatment does not induce changes in geographic student composition—a pattern consistent with features of India's engineering admissions system that limit post-admission sorting.

Together, institutional context and empirical evidence support our interpretation of the estimated effects as changes in institution-level outcomes, rather than artifacts of treatment-induced student reallocation.

6.5 Can Boom in Local Labor Markets be Driving the Results?

If AI centers were disproportionately established in districts experiencing contemporaneous growth in AI-intensive sectors, our estimates could reflect local labor-demand shocks rather than institutional investments. Several pieces of evidence suggest this is unlikely.

First, AI production and hiring in India do not necessarily occur in the same districts where prestigious engineering colleges are located. For example, one of the most prominent IITs is located in Kanpur, yet the district has a relatively low IT-oriented employment share compared with major technology hubs such as Bangalore and Hyderabad.[29] In practice, graduates from selective institutions commonly relocate to major hubs after graduation, implying that district-level labor-market conditions around the campus are only weakly informative about realized placement outcomes.[30]

[28] See Bagde et al. (2016) for evidence on centralized allocation, limited post-admission mobility, and how admission ranks map into institutional quality.

[29] Using the 2016 Consumer Pyramids Household Survey (CPHS) collected by the Centre for Monitoring Indian Economy, we find that only 0.6% of employed individuals in Kanpur hold a degree in computer applications or engineering, compared with 6.4% in Bangalore and 6.8% in Hyderabad.

[30] See Bagde et al. (2016) for evidence on spatial mismatch between training institutions and high-productivity labor markets.

Second, evidence from online vacancy data indicates that the national demand shock for AI skills begins only in 2016. Using archival job postings from 2010-2019, Copestake et al. (2023) show that India's share of vacancies requiring machine-learning or AI skills rises sharply only after 2016, consistent with the diffusion of open-source ML tools.[31] Because our identification compares trends between treated and never-treated institutions and includes year fixed effects, this common national demand shock is absorbed by time variation rather than confounding treatment assignment.

Third, when vacancy trends are aggregated to the district level, existing evidence finds no meaningful changes in broader labor-market outcomes. Copestake et al. (2023) report no detectable effects on total vacancies, district-level employment, or wages in either the short or medium run, suggesting that localized AI booms are limited in this period.[32]

6.6 World-Wide Increase in AI Jobs

A related concern is that our estimates could be confounded by the worldwide expansion of AI-related employment rather than by the establishment of AI centers themselves. Our empirical design addresses this concern in two ways. First, year fixed effects absorb shocks common to all institutions in a given year, including aggregate growth in AI-related jobs. Second, the staggered establishment of AI centers across institutions allows identification to come from relative timing rather than from global trends. A global increase in AI demand therefore does not threaten identification unless it induces differential district-level labor-demand shocks that simultaneously affect graduate outcomes and the timing of AI center adoption. As discussed in Section 6.5, existing evidence suggests that AI demand growth during our study period was not geographically concentrated in a manner likely to drive either adoption timing or placement outcomes.

To further address this concern empirically, we allow for flexible geographic heterogeneity in exposure to global AI demand. We incorporate two sets of geographic controls in our specifications.[33] First, we include state fixed effects in the BJS specification to absorb time-invariant differences across states. Second, we allow for district-specific labor-market trends by interacting 2017 district-level employment, labor-force participa-

[31] Copestake et al. (2023) distinguish between "AI-demanding" and "AI-producing" firms and show that the large increase in AI-skill demand in 2016-2019 reflects a technology shock rather than changes in job-posting composition.

[32] Copestake et al. (2023) also show that increases in AI vacancies at adopting establishments reduce non-AI hiring and wage offers within the same establishment, but these within-establishment effects net out when aggregated to the district level, yielding no detectable impacts on total labor demand or NSS/PLFS employment outcomes.

[33] SDID does not allow for additional fixed effects by design.

tion, and unemployment rates with year dummies.[34] In addition, we include a measure of pre-period change in local labor market for IT related employment—the growth in the share of employed professionals between 2012 and 2017—interacted with year dummies.[35] These interactions flexibly account for heterogeneous district-level labor-market trajectories that might otherwise correlate with AI center adoption. The results are reported in Appendix F.5. As shown, the estimated effects remain robust after including these controls. These patterns mitigate concerns that our estimates are driven by geographically concentrated labor-demand shocks rather than by changes in institutional educational capital.

Overall, the above robustness checks show that the estimated effects are not artifacts of sample selection, elite-institution trends, or local labor-market booms. The results remain directionally stable and economically meaningful under alternative samples, alternative timing windows, and stricter geographic controls, reinforcing the credibility of the causal interpretation.

7 Discussion of Limitations

Our findings should be interpreted with several caveats that follow from the scope of our data, which are observed at the institution-year level rather than at the level of students, courses, or firms. First, we study the upper tier of India's engineering ecosystem, where baseline resources and academic preparedness are relatively strong; institutions facing more binding resource constraints may experience different impacts, which remains an open question. Second, although we document improvements in employment and earnings, our outcome measures capture only the first placement following graduation, and we cannot observe longer-run outcomes such as job mobility, career progression, or productivity as exposed cohorts progress through the labor market. Third, our evidence on mechanisms relies on institutional expenditures and externally funded research projects, and we cannot observe the specific content of AI-focused curricula, the degree of student participation in AI center activities, or firm-side adjustments in recruitment strate-

[34]The district-level labor-market measures are drawn from the Periodic Labor Force Survey (PLFS). We use 2017 values because PLFS was first launched in 2017-18. The three variables—workforce participation rate (share of population in the labor force), worker-population ratio (share of the population employed), and unemployment rate (share of the labor force unemployed)—capture baseline cross-district differences in labor-market depth and employment opportunities.

[35]Using the NSS 68th Round Employment-Unemployment Survey (2011-12) and the PLFS 2017-18, we compute the share of employed individuals classified as professionals, technicians, and associate professionals (National Classification of Occupations-2004) in each district. We then construct the growth measure as the proportional change in this share between 2012 and 2017.

gies. As a result, we cannot separately distinguish skill accumulation from changes in matching, employer composition, or selection into placement reporting. Future work linking student-level coursework records, placement microdata, or employer demand signals would help illuminate these channels more precisely. Despite these limitations, the results provide credible institution-level evidence that targeted, technology-aligned investments in higher education can shape early-career labor-market outcomes in the AI transition.

8 Conclusion

The rapid diffusion of AI has heightened the importance of AI-complementary human capital for labor-market success. More broadly, it underscores a central question in the economics of frontier technology: when do new technological opportunities translate into broad labor-market gains, and when are those gains limited by institutional capacity? This paper provides causal evidence on that question by studying how the establishment of AI centers within higher-education institutions affects student outcomes. Using institution-by-year administrative data and recent staggered-adoption difference-in-differences estimators, we find sizable and persistent increases in full-time placement and early-career earnings for graduates of four-year programs.

The heterogeneity across institutions aligns closely with our conceptual framework. On the extensive margin, placement gains are concentrated among institutions serving academically stronger cohorts, consistent with complementarities between AI-center investments and baseline preparedness that relax a placement barrier. On the intensive margin, conditional earnings gains vary systematically with cohorts' counterfactual opportunities and market access. Institutions historically serving cohorts with weaker outside options exhibit larger proportional earnings gains among placed graduates, while institutions serving more geographically mobile cohorts experience larger conditional earnings increases, consistent with improved access to high-return AI labor markets once graduates are placed. We also find that treated institutions expand the inputs most directly related to AI-oriented skill formation, including computational laboratories, specialized faculty, and externally funded research. Taken together, these patterns support the interpretation that AI centers relax placement and wage-returns barriers rather than merely track pre-existing trends.

As countries confront the skill demands of the AI transition, understanding how higher-education systems can build AI-complementary human capital, and how such investments influence the distribution of opportunity, remains a central economic and policy

challenge.

References

Acemoglu, Daron, "Why do new technologies complement skills? Directed technical change and wage inequality," *The quarterly journal of economics*, 1998, *113* (4), 1055–1089.

_ **and David Autor**, "Skills, tasks and technologies: Implications for employment and earnings," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 1043–1171.

_ **and Pascual Restrepo**, "Artificial intelligence, automation, and work," in "The economics of artificial intelligence: An agenda," University of Chicago Press, 2018, pp. 197–236.

_ **, David Autor, Jonathon Hazell, and Pascual Restrepo**, "Artificial intelligence and jobs: Evidence from online vacancies," *Journal of Labor Economics*, 2022, *40* (S1), S293–S340.

_ **, Simon Johnson, and James A Robinson**, "Institutions as a fundamental cause of long-run growth," *Handbook of economic growth*, 2005, *1*, 385–472.

Arkhangelsky, Dmitry, Susan Athey, David A Hirshberg, Guido W Imbens, and Stefan Wager, "Synthetic difference-in-differences," *American Economic Review*, 2021, *111* (12), 4088–4118.

Bagde, Surendrakumar, Dennis Epple, and Lowell Taylor, "Does affirmative action work? Caste, gender, college quality, and academic success in India," *American Economic Review*, 2016, *106* (6), 1495–1521.

Bloom, Nicholas, Mirko Draca, and John Van Reenen, "Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity," *The review of economic studies*, 2016, *83* (1), 87–117.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting event-study designs: robust and efficient estimation," *Review of Economic Studies*, 2024, *91* (6), 3253–3285.

Bryan, Gharad T, Kyra Frye, and Melanie Morten, "Spatial economics for low-and middle-income countries," Technical Report, National Bureau of Economic Research 2025.

Brynjolfsson, Erik, Bharat Chandar, and Ruyu Chen, "Canaries in the coal mine? six facts about the recent employment effects of artificial intelligence," 2025.

_ **, Danielle Li, and Lindsey Raymond**, "Generative AI at work," *The Quarterly Journal of Economics*, 2025, *140* (2), 889–942.

Callaway, Brantly and Pedro HC Sant'Anna, "Difference-in-differences with multiple time periods," *Journal of econometrics*, 2021, *225* (2), 200–230.

Clarke, Damian, Daniel Pailanir, Susan Athey, and Guido Imbens, "On synthetic difference-in-differences and related estimation methods in Stata," *The Stata Journal*, 2024, *24* (4), 557–598.

Cohen, Wesley M and Daniel A Levinthal, "Innovation and learning: the two faces of R & D," *The economic journal*, 1989, *99* (397), 569–596.

Copestake, Alexander, Max Marczinek, Ashley Pople, and Katherine Stapleton, "Ai and services-led growth: Evidence from indian job adverts," Technical Report, Working Paper 2023.

Demirci, Ozge, Jonas Hannane, and Xinrong Zhu, "Who is AI replacing? The impact of generative AI on online freelancing platforms," *Management Science*, 2025.

Demombynes, Gabriel, Jörg Langbein, and Michael Weber, *The Exposure of Workers to Artificial Intelligence in Low-and Middle-Income Countries*, World Bank, 2025.

Desai, Sonalde, Elizabeth Adams, and Amaresh Dubey, "Human Development in India: Challenges for a Society in Transition," *Oxford University Press*, 2010.

Goldin, Claudia and Lawrence F Katz, "The race between education and technology," in "Inequality in the 21st Century," Routledge, 2018, pp. 49–54.

Hausmann, Ricardo, Jason Hwang, and Dani Rodrik, "What you export matters," *Journal of Economic Growth*, 2007.

Heckman, James J. and Bo Honoré, "The Empirical Content of the Roy Model," *Econometrica*, 1990, *58* (5), 1121–1149.

Hnatkovska, Viktoria, Amartya Lahiri, and Sourabh Paul, "Castes and labor mobility," *American Economic Journal: Applied Economics*, 2012, *4* (2), 274–307.

_ and _, "The rural-urban divide in India," *International Growth Centre Working Paper*, 2013.

Hui, Xiang, Oren Reshef, and Luofeng Zhou, "The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market," *Organization Science*, 2024, *35* (6), 1977–1989.

Kahn, Shulamit and Donna Ginther, "Women and STEM," Technical Report, National Bureau of Economic Research 2017.

Krueger, Alan B., "How Computers Have Changed the Wage Structure: Evidence from Microdata," *Quarterly Journal of Economics*, 1993, *108* (1), 33–60.

Mani, Deepa, Saurabh Tomar, Nikhil Madan, and Anshul Bhatia, "The Impact of AI on the Indian Labour Market," Technical Report, Intel 2020. Industry report.

Maslej, Nestor, Loredana Fattorini, Raymond Perrault, Yolanda Gil, Vanessa Parli, Njenga Kariuki, Emily Capstick, Anka Reuel, Erik Brynjolfsson, John Etchemendy et al., "Artificial intelligence index report 2025," *arXiv preprint arXiv:2504.07139*, 2025.

Muralidharan, Karthik and Nishith Prakash, "Cycling to School: Increasing Secondary School Enrollment for Girls in India," *American Economic Journal: Applied Economics*, 2017, *9* (3), 321–350.

Nelson, Richard R, *National Innovation Systems: A Comparative Analysis*, Oxford University Press, 1993.

Ravindran, B., S. Sarawagi, and A. Jain, "AI and Data Science Research Centers in Indian Academic Institutions," *Communications of the ACM* 2022. Accessed October 2025.

Roy, A. D., "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 1951, *3* (2), 135–146.

Sahoo, Soham and Stephan Klasen, "Gender segregation in education: Evidence from higher secondary stream choice in India," *Demography*, 2021, *58* (3), 987–1010.

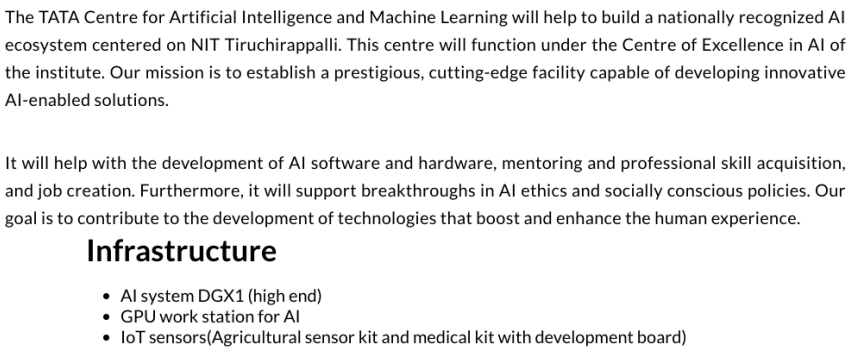
Webb, Michael, "The Impact of Artificial Intelligence on the Labor Market," *Working Paper*, 2020. Stanford University.

World Bank, "South Asia Development Update, October 2025: Jobs, AI, and Trade," Technical Report, World Bank 2025. Accessed October 2025.

Yotzov, Ivan, Jose Maria Barrero, Nicholas Bloom, Philip Bunn, Steven J Davis, Kevin M Foster, Aaron Jalca, Brent H Meyer, Paul Mizen, Michael A Navarrete et al., "Firm Data on AI," Technical Report, National Bureau of Economic Research 2026.

Figures

FIGURE 1: Illustrative Examples of AI-Center Educational Capital



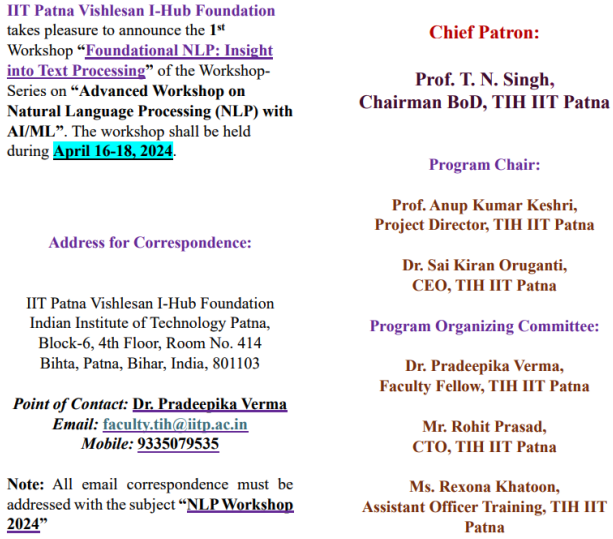
The TATA Centre for Artificial Intelligence and Machine Learning will help to build a nationally recognized AI ecosystem centered on NIT Tiruchirappalli. This centre will function under the Centre of Excellence in AI of the institute. Our mission is to establish a prestigious, cutting-edge facility capable of developing innovative AI-enabled solutions.

It will help with the development of AI software and hardware, mentoring and professional skill acquisition, and job creation. Furthermore, it will support breakthroughs in AI ethics and socially conscious policies. Our goal is to contribute to the development of technologies that boost and enhance the human experience.

Infrastructure

- AI system DGX1 (high end)
- GPU work station for AI
- IoT sensors(Agricultural sensor kit and medical kit with development board)

(a) AI Infrastructure



IIT Patna Vishlesan I-Hub Foundation takes pleasure to announce the **1st** Workshop **"Foundational NLP: Insight into Text Processing"** of the Workshop-Series on **"Advanced Workshop on Natural Language Processing (NLP) with AI/ML"**. The workshop shall be held during **April 16-18, 2024**.

Address for Correspondence:

IIT Patna Vishlesan I-Hub Foundation
Indian Institute of Technology Patna,
Block-6, 4th Floor, Room No. 414
Bihta, Patna, Bihar, India, 801103

Point of Contact: **Dr. Pradeepika Verma**
Email: faculty.tih@iitp.ac.in
Mobile: **9335079535**

Note: All email correspondence must be addressed with the subject **"NLP Workshop 2024"**

Chief Patron:

Prof. T. N. Singh,
Chairman BoD, TIH IIT Patna

Program Chair:

Prof. Anup Kumar Keshri,
Project Director, TIH IIT Patna

Dr. Sai Kiran Oruganti,
CEO, TIH IIT Patna

Program Organizing Committee:

Dr. Pradeepika Verma,
Faculty Fellow, TIH IIT Patna

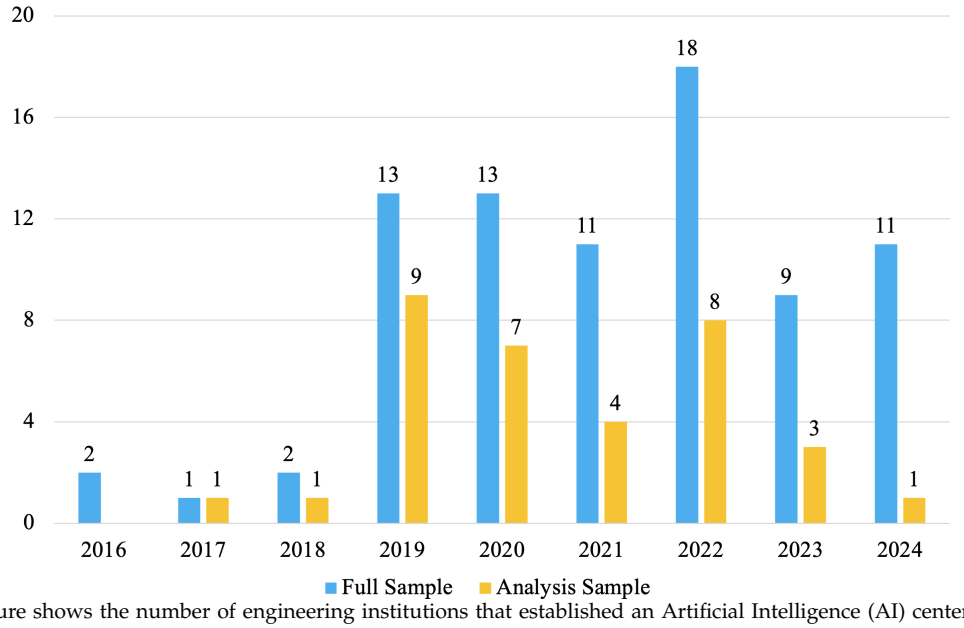
Mr. Rohit Prasad,
CTO, TIH IIT Patna

Ms. Rexona Khatoon,
Assistant Officer Training, TIH IIT Patna

(b) AI Workshop Series

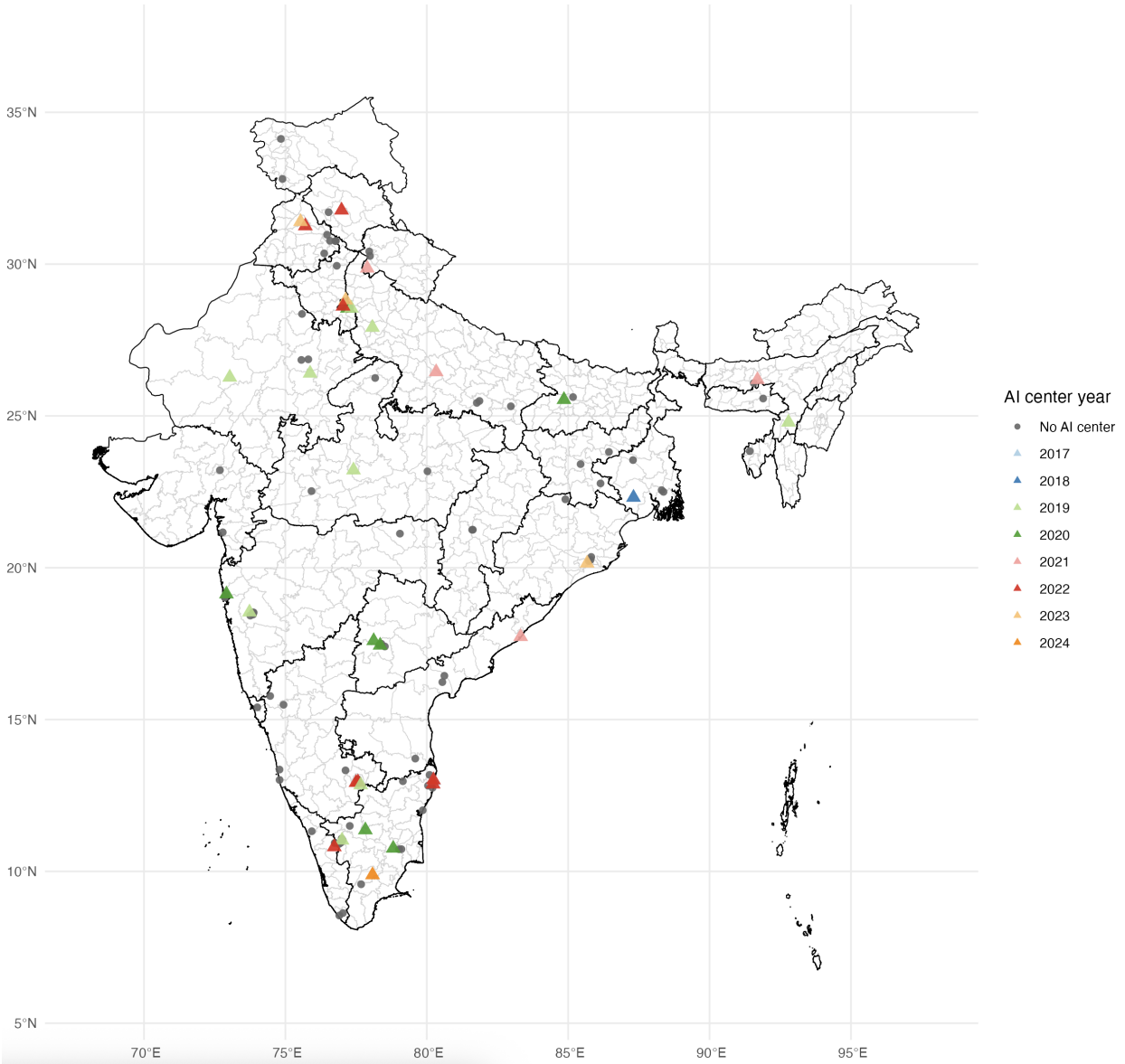
Notes: This figure presents two examples of activities associated with institution-level AI centers. Panel (a) shows AI-related computing infrastructure at National Institute of Technology Tiruchirappalli's TATA Centre for Artificial Intelligence and Machine Learning, including high-performance hardware and related technical resources. Panel (b) shows Indian Institute of Technology Patna Vishlesan I-Hub Foundation's 2024 workshop announcement for "Foundational NLP: Insight into Text Processing," illustrating specialized AI/ML training activity.

FIGURE 2: Distribution of AI Center Establishment Years Among Engineering Institutions



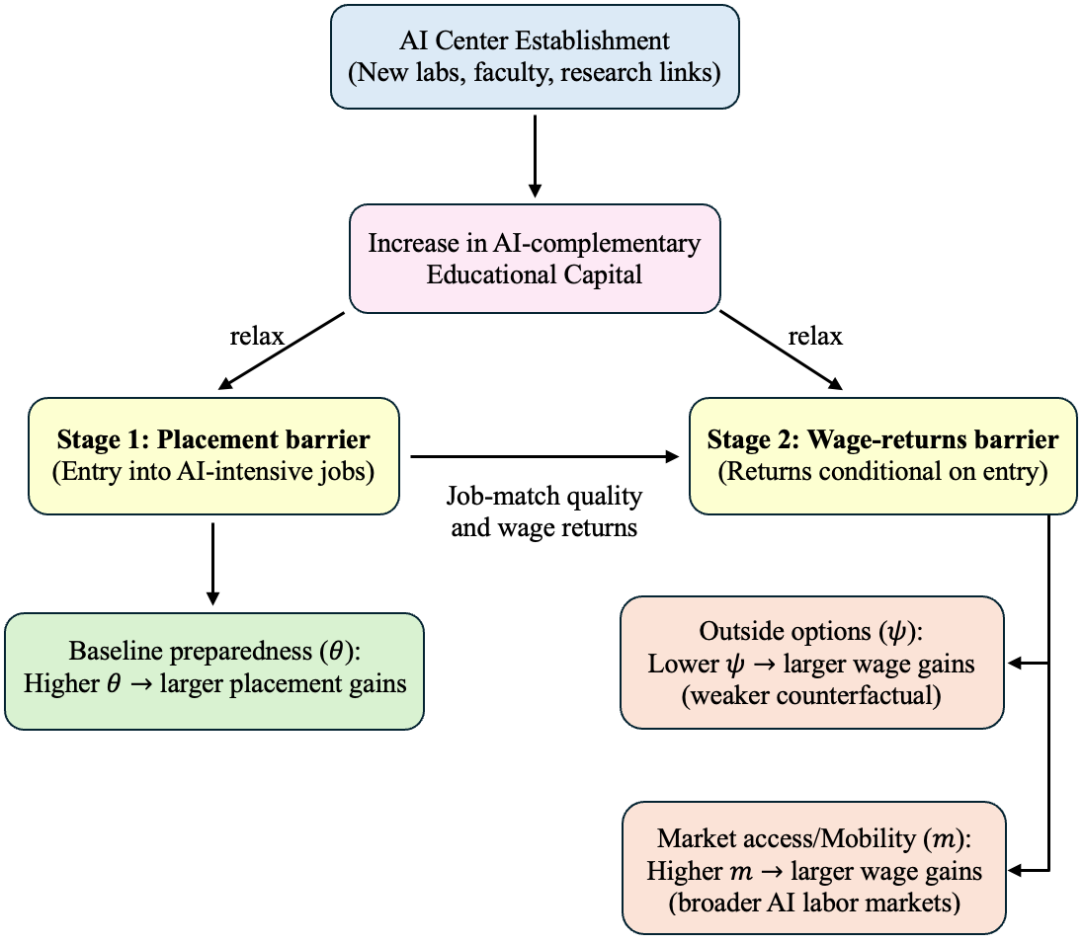
Notes: This figure shows the number of engineering institutions that established an Artificial Intelligence (AI) center between 2016 and 2024. We examined 268 engineering institutions in India and identified 87 with an AI center, of which seven lacked establishment years and were excluded from the figure. Among the remaining 80 institutions with confirmed establishment years, the number of new AI centers increased sharply after 2018, peaking in 2022 when 18 institutions established centers. The data are based on manual verification of institutional websites, social media pages, and major news outlets. The blue bars represent the 80 institutions with an AI center established between 2016 and 2024 among the 268 institutions, while the yellow bars represent the 34 elite institutions with an average NIRF engineering rank of 100 or better that established an AI center between 2017 and 2024.

FIGURE 3: Map of 113 Engineering Institutions



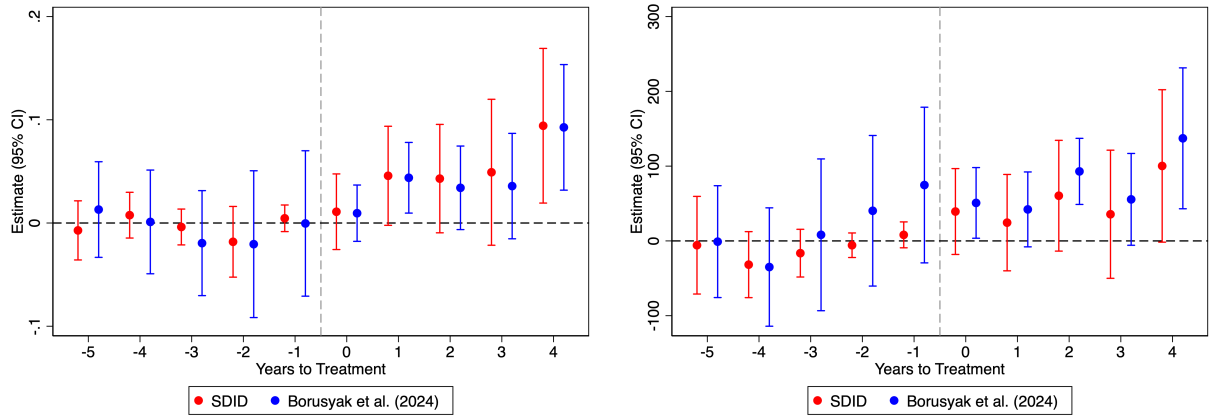
Notes: This figure shows the geographic distribution of the 113 engineering institutions included in the main analysis. Triangles denote institutions that established an AI center, with colors indicating the year of establishment between 2017 and 2024. Institutions without an AI center as of 2024 are shown as gray circles.

FIGURE 4: Conceptual Framework: Placement and Wage-returns Barriers



Notes: The framework illustrates two channels through which AI-center establishment affects labor-market outcomes. First, a *placement barrier*: AI-center investments raise AI-complementary skills, increasing the efficiency with which baseline preparedness translates into entry into AI-intensive jobs. Second, a *wage-returns barrier*: conditional on entry, wage gains depend on both counterfactual outside options and market access. Graduates with weaker outside options experience larger proportional gains, and cohorts with greater geographic mobility access higher-return AI labor markets and better matches, amplifying conditional earnings gains. The framework operates at the institution level and does not rely on individual-level sorting.

FIGURE 5: Dynamic Effects on Four-Year Undergraduate Program Outcomes



(a) % Placed

(b) Salary (1,000 INR)

Notes: We employ two estimation methods on the institution-year panel data (2016-2024). Each panel reports event study estimates of the impacts of AI Center establishment on four-year undergraduate outcomes. "% Place" = The proportion of graduates from four-year undergraduate programs who obtained employment. "Salary (1,000 INR)" = The median salary of employed graduates from four-year undergraduate programs, expressed in thousands of Indian Rupees. All monetary variables are converted to real terms using the CPI (2016 = 100). Red estimates are obtained using Synthetic Difference-in-Differences and blue estimates are obtained using the Borusyak-Jaravel-Spiess imputation estimator (Borusyak et al., 2024). The horizontal black dashed line denotes zero. The point at Years to Treatment = 0 corresponds to the year in which an AI center was established. Confidence intervals are at the 95 percent.

Tables

TABLE 1: Descriptive Statistics of Outcome Variables (2016-2018)

Variable	Obs	Mean	Std. dev.
Four-Year Undergraduate Programs			
% Higher Studies	286	0.117	0.098
% Placed	286	0.706	0.138
Salary (1,000 INR)	286	585.362	280.632
Two-Year Master's Programs			
% Higher Studies	296	0.118	0.112
% Placed	296	0.484	0.272
Salary (1,000 INR)	294	556.165	305.304
Capital Expenditure (million INR)			
Library	298	26.961	35.867
Equipment	298	113.993	194.092
Eng. Workshops	298	17.130	42.234
Other Cap.	298	87.501	194.527
Operating Expenditure (million INR)			
Salaries	298	581.588	596.503
Maintenance	298	408.401	551.599
Seminars	298	16.168	41.161
Sponsored Research Projects (2018)			
# of Projects	101	93.119	154.180
# of Agencies	101	26.426	36.882
Total Amount (million INR)	101	225.522	652.497
Consultancy Projects (2018)			
# of Projects	101	140.733	274.712
# of Agencies	101	239.495	409.759
Total Amount (million INR)	101	55.444	125.218

Notes: We restrict the sample to institutions with an average NIRF engineering rank of 100 or better. "% Higher Studies" = The proportion of graduates from four-year undergraduate programs (or two-year master's programs) who pursue higher studies. "% Place" = The proportion of graduates from four-year undergraduate programs (or two-year master's programs) who obtained employment. "Salary (1,000 INR)" = The median salary of employed graduates from four-year undergraduate programs (or two-year master's programs), expressed in thousands of Indian Rupees (1,000 INR). All monetary variables are converted to real terms using the CPI (2016 = 100). "Library" = The annual capital expenditure on library including books, journals, and e-Resources, expressed in millions of Indian Rupees. "Equipment" = The annual capital expenditure on new equipment and software for laboratories, expressed in millions of Indian Rupees. "Eng. Workshops" = The annual capital expenditure on establishing or upgrading engineering workshops, including machinery and tools, expressed in millions of Indian Rupees. "Other Cap." = The annual capital expenditure on the creation of other capital assets, such as setting up classrooms, seminar halls, conference halls, libraries, laboratories, and engineering workshops, excluding expenditure on land and buildings, expressed in millions of Indian Rupees. "Salaries" = The annual operating expenditure on salaries for teaching and non-teaching staff, expressed in millions of Indian Rupees. "Maintenance" = The annual operating expenditure on the maintenance of academic infrastructure, consumables, and other running expenses (excluding maintenance of hostels and allied services, building rent, depreciation costs, etc.), expressed in millions of Indian Rupees. "Seminar" = The annual operating expenditure on seminars, conferences, and workshops, expressed in millions of Indian Rupees.

TABLE 2: Effects of AI Center Establishment on Graduate Outcomes: Four-Year Undergraduate Programs (2016-2024)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	% Higher Studies		% Placed		Salary (1,000 INR)	
	SDID	BJS	SDID	BJS	SDID	BJS
Treatment	-0.007	-0.010	0.043**	0.038**	51.613*	66.455***
	(0.016)	(0.011)	(0.016)	(0.020)	(28.744)	(21.975)
Placebo SE	[0.015]		[0.026]*		[26.406]*	
2016 Mean	0.11	0.10	0.73	0.74	604.90	576.87
No. of Institutions	72	107	72	107	72	107
N	648	857	648	857	648	857

*Notes:*We employ two estimation methods on the institution-year panel (2016–2024). Columns (1), (3), and (5) report estimates from the synthetic difference-in-differences (SDID) method with staggered adoption (Arkhangelsky et al., 2021). Columns (2), (4), and (6) report estimates from the DID imputation method (Borusyak et al., 2024). Standard errors, shown in parentheses, are obtained via bootstrap inference with 999 replications in columns (1), (3), and (5), and are clustered at the institution level in columns (2), (4), and (6). As a robustness check, we additionally report placebo-based standard errors with 999 replications for columns (1), (3), and (5). "% Higher Studies" = The proportion of graduates from four-year undergraduate programs who pursue higher studies. "% Place" = The proportion of graduates from four-year undergraduate programs who obtained employment. "Salary (1,000 INR)" = The median salary of employed graduates from four-year undergraduate programs, expressed in thousands of Indian Rupees (1,000 INR). All monetary variables are converted to real terms using the CPI (2016 = 100). "2016 Mean" is the average value of the respective variable in 2016. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 3: Effects of AI Center Establishment on Graduate Outcomes: Two-Year Master's Programs (2016-2024)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	% Higher Studies		% Placed		Salary (1,000 INR)	
	SDID	BJS	SDID	BJS	SDID	BJS
Treatment	-0.001	0.014	-0.048	-0.030	9.710	44.568
	(0.022)	(0.022)	(0.036)	(0.030)	(45.379)	(32.353)
Placebo SE	[0.024]		[0.042]		[41.091]	
2016 Mean	0.11	0.10	0.47	0.48	573.08	552.55
No. of Institutions	73	112	73	112	73	112
N	657	883	657	883	657	883

Notes: We employ two estimation methods on the institution-year panel (2016–2024). Columns (1), (3), and (5) report estimates from the synthetic difference-in-differences (SDID) method with staggered adoption (Arkhangelsky et al., 2021). Columns (2), (4), and (6) report estimates from the DID imputation method (Borusyak et al., 2024). Standard errors, shown in parentheses, are obtained via bootstrap inference with 999 replications in columns (1), (3), and (5), and are clustered at the institution level in columns (2), (4), and (6). As a robustness check, we additionally report placebo-based standard errors with 999 replications for columns (1), (3), and (5). "% Higher Studies" = The proportion of graduates from two-year master's programs who pursue higher studies. "% Place" = The proportion of graduates from two-year master's programs who obtained employment. "Salary (1,000 INR)" = The median salary of employed graduates from two-year master's programs, expressed in thousands of Indian Rupees (1,000 INR). All monetary variables are converted to real terms using the CPI (2016 = 100). "2016 Mean" is the average value of the respective variable in 2016. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 4: Heterogeneous Effects of AI Center Establishment on Graduate Outcomes: Four-Year Undergraduate Programs (2016-2024)

	(1)	(2)	(3)	(4)
	Share of Socially Challenged Students		Share of Within State Students	
	SDID	BJS	SDID	BJS
Panel A. % Placed				
Treatment (Above)	0.038	0.016	0.029	0.024
	(0.031)	(0.027)	(0.042)	(0.038)
Treatment (Below)	0.086**	0.081***	0.050*	0.041**
	(0.037)	(0.016)	(0.029)	(0.016)
Difference	-0.047	-0.065**	-0.020	-0.017
	(0.048)	(0.028)	(0.051)	(0.039)
Panel B. Salary (1,000 INR)				
Treatment (Above)	41.953	94.804***	10.860	-47.060
	(40.521)	(29.030)	(49.877)	(48.216)
Treatment (Below)	26.104	2.200	56.266	103.878***
	(47.214)	(32.003)	(38.578)	(23.391)
Difference	15.848	92.605**	-45.406	-150.938***
	(62.218)	(38.432)	(63.055)	(49.330)
2016 Mean (% Placed)	0.73	0.74	0.73	0.74
2016 Mean (Salary)	607.08	576.87	607.08	576.87
No. of Institutions	72	94	72	94
N	648	777	648	777

Notes: We employ two estimation methods on the institution-year panel (2016–2024). Columns (1) and (3) report estimates from the synthetic difference-in-differences (SDID) method with staggered adoption (Arkhangelsky et al., 2021). Columns (2) and (4) report estimates from the DID imputation method (Borusyak et al., 2024). Standard errors, shown in parentheses, are obtained via bootstrap inference with 999 replications in columns (1) and (3), and are clustered at the institution level in columns (2) and (4). "Treatment (Above)" reports estimates for the subsample of institutions with an average share of socially challenged students above the median in columns (1)-(2), and for the subsample of institutions with an average share of within-state students above the median in columns (3)-(4). "Treatment (Below)" reports estimates for the subsample of institutions with an average share of socially challenged students below the median in columns (1)-(2), and for the subsample of institutions with an average share of within-state students below the median in columns (3)-(4). "% Place" = The proportion of graduates from four-year undergraduate programs who obtained employment. "Salary (1,000 INR)" = The median salary of employed graduates from four-year undergraduate programs, expressed in thousands of Indian Rupees (1,000 INR). All monetary variables are converted to real terms using the CPI (2016 = 100). "2016 Mean" is the average value of the respective variable in 2016. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 5: Impact of AI Center Establishment on Capital & Operating Expenditures (2016-2024)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Capital Expenditures			Operating Expenditures		
VARIABLES	Library	Equipment	Eng. Workshops	Other Cap.	Salaries	Maintenance	Seminars
Panel A. SDID							
Treatment	0.164	135.078*	-22.364	17.049	266.995**	143.858	-3.279
Bootstrapped SE	(3.497)	(79.979)	(20.523)	(27.072)	(125.632)	(149.410)	(8.532)
Placebo SE	[3.272]	[24.989]***	[7.496]***	[45.119]	[67.229]***	[89.523]	[5.878]
2016 Mean	28.93	132.87	22.69	93.04	630.09	455.64	15.51
No. of Institutions	75	75	75	75	75	75	75
N	675	675	675	675	675	675	675
Panel B. BJS							
Treatment	-0.766	91.324***	-24.972***	16.805	216.798***	202.209***	0.954
	(2.216)	(33.969)	(5.018)	(23.684)	(60.786)	(65.199)	(4.948)
2016 Mean	23.84	109.98	18.78	78.23	551.41	378.77	12.95
No. of Institutions	110	110	110	110	110	110	110
N	893	893	893	893	893	893	893

Notes: We employ two estimation methods on the institution-year panel (2018–2024). Panel A reports estimates from the synthetic difference-in-differences (SDID) method with staggered adoption (Arkhangelsky et al., 2021). Panel B reports estimates from the DID imputation method (Borusyak et al., 2024). Standard errors, reported in parentheses, are obtained via bootstrap inference with 999 replications in Panel A, and are clustered at the institution level in Panel B. As a robustness check, Panel A also reports placebo-based standard errors computed from 999 replications. "Library" = The annual capital expenditure on library including books, journals, and e-Resources. "Equipment" = The annual capital expenditure on new equipment and software for laboratories. "Eng. Workshops" = The annual capital expenditure on establishing or upgrading engineering workshops, including machinery and tools. "Other Cap." = The annual capital expenditure on the creation of other capital assets, such as setting up classrooms, seminar halls, conference halls, libraries, laboratories, and engineering workshops, excluding expenditure on land and buildings. "Salaries" = The annual operating expenditure on salaries for teaching and non-teaching staff. "Maintenance" = The annual operating expenditure on the maintenance of academic infrastructure, consumables, and other running expenses (excluding maintenance of hostels and allied services, building rent, depreciation costs, etc.). "Seminar" = The annual operating expenditure on organizing seminars, conferences, and workshops, including event logistics, materials, and related recurring costs. The units of all outcome variables are in million INR. All monetary variables are converted to real terms using the CPI (2016 = 100). "2016 Mean" is the average value of the respective variable in 2016. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 6: Impact of AI Center Establishment on Research Projects (2018-2024)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Sponsored Research			Consultancy Projects		
	# of Projects	# of Agencies	Total Amount (million INR)	# of Projects	# of Agencies	Total Amount (million INR)
Panel A. SDID						
Treatment	40.501	11.351	90.304*	-13.240	-85.910	32.217
Bootstrapped SE	(30.561)	(9.924)	(51.820)	(65.475)	(107.901)	(27.236)
Placebo SE	[18.977]**	[6.309]*	[20.345]***	[76.908]	[163.981]	[10.314]***
2018 Mean	98.86	27.71	220.66	135.51	234.94	52.69
No. of Institutions	78	78	78	78	78	78
N	546	546	546	546	546	546
Panel B. ***BJS***						
Treatment	36.810	10.960	59.612	-28.395	-80.479	17.421*
	(25.600)	(8.192)	(39.390)	(53.357)	(99.852)	(10.374)
2018 Mean	90.50	26.32	204.78	134.70	233.97	49.92
No. of Institutions	106	106	106	106	106	106
N	682	682	682	682	682	682

Notes: We employ two estimation methods on the institution-year panel (2018–2024). Panel A reports estimates from the synthetic difference-in-differences (SDID) method with staggered adoption (Arkhangelsky et al., 2021). Panel B reports estimates from the DID imputation method (Borusyak et al., 2024). Standard errors, reported in parentheses, are obtained via bootstrap inference with 999 replications in Panel A, and are clustered at the institution level in Panel B. As a robustness check, Panel A also reports placebo-based standard errors computed from 999 replications. Columns (1)-(3) report the results for sponsored research projects, while columns (4)-(6) present the results for consultancy projects. "# of Projects" = The total number of sponsored projects at the institution level in column (1) and the total number of consultancy projects in column (4). "# of Agencies" = The total number of funding agencies in column (2) and the total number of client organizations in column (5). "Total Amount (million INR)" = The total funding received for sponsored research projects in column (3) and the total amount received for consultancy projects in column (6). The units of outcome variables in columns 3 and 6 are in million INR. All monetary variables are converted to real terms using the CPI (2016 = 100). "2018 Mean" is the average value of the respective variable in 2018. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 7: Impact of AI Center Establishment on AI-related Publications (2016-2024)

VARIABLES	(1)	(2)	(3)	(4)
	Top-100		Full Sample	
	SDID	BJS	SDID	BJS
Treatment	6.970	12.932*	9.035	13.020**
	(8.174)	(7.694)	(5.733)	(5.260)
Placebo SE	[8.368]		[4.432]**	
2016 Mean	8.71	8.49	5.12	4.92
No. of Institutions	96	112	205	262
N	864	975	1,845	2,212

Notes: We employ two estimation methods on the institution-year panel (2016–2024). Columns (1) and (3) report estimates from the synthetic difference-in-differences (SDID) method with staggered adoption (Arkhangelsky et al., 2021). Columns (2) and (4) report estimates from the DID imputation method (Borusyak et al., 2024). Standard errors, shown in parentheses, are obtained via bootstrap inference with 999 replications in columns (1) and (3), and are clustered at the institution level in columns (2) and (4). As a robustness check, we additionally report placebo-based standard errors with 999 replications for columns (3) and (6). The dependent variable is the number of AI-related publications. Columns (1)–(3) use the main analysis sample, consisting of institutions with an average rank of 100 or better, while columns (4)–(6) use the full sample. "2016 Mean" is the average value of the respective variable in 2016. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Online Appendix (For Online Publication Only)

Investing in AI Talent: Labor-Market Returns to AI Skilling College Infrastructure in India

Yutong Chen Sheetal Sekhri

A Appendix Roadmap

This Online Appendix provides additional institutional background, details on data construction, pre-trend diagnostics, heterogeneity tables, and robustness checks including placebo re-timings, tests for differential macro trend exposure, and alternative samples, following standard guidance for staggered-adoption designs.

Appendix B provides extended policy context and summarizes the roll-out of AI centers. Appendix C details the NIRF data cleaning procedures, variable construction, and sample restrictions. Appendix D provides additional information on the implementation of the SDID estimator. Appendix E reports additional results. Appendix F summarizes robustness checks, including alternative samples and specifications with additional controls. Appendix G discusses the conceptual framework.

B Additional Policy Background

In February 2018, the Hon'ble Finance Minister, Arun Jaitley, highlighted artificial intelligence (AI) in the Union Budget speech for the first time, emphasizing its importance for advancing India's digital economy and strengthening the Digital India agenda.

> "108. Global economy is transforming into a digital economy thanks to development of cutting edge technologies in digital space - machine learning, artificial intelligence, internet of things, 3D printing and the like. Initiatives such as Digital India, Start Up India, Make in India would help India establish itself as a knowledge and digital society. NITI Aayog will initiate a national program to direct our efforts in the area of artificial intelligence, including research and development of its applications."

This directive assigned the National Institution for Transforming India (NITI Aayog) responsibility for developing a national program on AI. In response, NITI Aayog released

the National Strategy for Artificial Intelligence (*#AIforAll*) on June 4, 2018, outlining India's AI vision and implementation roadmap.

The strategy identified three focus areas: research, skilling, and adoption, and called for strong institutional infrastructure to accelerate AI development. A major recommendation was the establishment of a network of Centers of Excellence (CoEs) with dedicated funding on the order of INR 0.5∼1 billion per center. These centers were envisioned to be hosted at India's premier academic institutions, including the Indian Institute of Science (IISc), Indian Statistical Institute (ISI), and leading Indian Institutes of Technology (IITs) and Indian Institutes of Information Technology (IIITs). By bringing together skilled faculty, research infrastructure, and industry collaboration, such centers could potentially enhance educational quality and strengthen the pipeline of graduates equipped for emerging technical fields.

However, implementation of this strategy progressed gradually. It was not until 2023 that the Ministry of Education launched a coordinated initiative to establish AI CoEs through a nationwide call for proposals. The scheme focuses on three domains central to India's development—agriculture, health, and sustainable cities—and explicitly aims to "Make AI in India and Make AI Work for India." The lead institute of each proposal was required to be ranked within the top 100 of the National Institutional Ranking Framework (NIRF). The program committed INR 9.9 billion over financial years 2023-24 to 2027-28. From 55 submissions, three proposals were selected in early September 2024, involving 11 premier institutions, such as IIT Delhi, IIT Kanpur, and IIT Ropar.

Meanwhile, India's leading academic institutions have proactively created AI-oriented centres funded through combinations of alumni philanthropy, industry partnership, and local government support. For example, IIT Kharagpur established the Centre for AI in 2018 (seed-funded by Capillary Technologies); Andhra University, Visakhapatnam campus, launched the CoE for Internet of Things and AI in 2021, jointly funded by the Ministry of Electronics & Information Technology, the Government of Andhra Pradesh, and the National Association of Software & Services Companies; and G.L. Bajaj Institute of Technology and Management established the NVIDIA AI Learning and Research Center in 2022 (funded by NVIDIA). These examples shows the diverse funding models and institutional pathways through which AI-focused centers are emerging across the country. We display the list of institutions with AI center establishment year in Table B8.

TABLE B8: Institutions by AI Center Establishment Year

Year	Institution	Avg. Rank
2016	Indraprastha Institute of Information Technology Delhi	65.3
2016	Vels Institute of Science, Technology & Advanced Studies (VISTAS)	142.0
2017	Indian Institute of Technology Madras	1.0
2018	Indian Institute of Technology Kharagpur	4.7
2018	Nitte Meenakshi Institute of Technology	144.2
2019	Aligarh Muslim University	35.6
2019	Amity University	33.4
2019	Banasthali Vidyapith	68.6
2019	Chitkara University	118.8
2019	Dayalbagh Educational Institute	100.2
2019	Government Engineering College, Thrissur	160.0
2019	Indian Institute of Technology Jodhpur	42.2
2019	Institute of Aeronautical Engineering	156.0
2019	International Institute of Information Technology Bangalore	69.6
2019	Maulana Azad National Institute of Technology	67.3
2019	National Institute of Technology Silchar	47.0
2019	PSG College of Technology	51.3
2019	Symbiosis International University	46.0
2020	Chennai Institute of Technology	175.0
2020	Dhirubhai Ambani Institute of Information and Communication Technology	102.6
2020	Indian Institute of Technology Bombay	2.8
2020	Indian Institute of Technology Delhi	2.3
2020	Indian Institute of Technology Hyderabad	8.3
2020	Indian Institute of Technology Patna	26.6
2020	International Institute of Information Technology Hyderabad	50.1
2020	K.S. Rangasamy College of Technology	99.0
2020	Kakatiya Institute of Technology & Science	188.5
2020	National Institute of Technology Tiruchirappalli	9.4
2020	Nirma University	124.8
2020	Rajiv Gandhi Institute of Petroleum Technology	120.0
2020	SR University	112.0
2021	AU College of Engineering	79.3
2021	Christ University	119.3
2021	G. L. A. University	186.0
2021	Indian Institute of Technology Guwahati	7.1
2021	Indian Institute of Technology Kanpur	4.3
2021	Indian Institute of Technology Roorkee	5.9
2021	Indira Gandhi Delhi Technical University for Women	161.3
2021	MIT World Peace University	139.5
2021	SVKM's Narsee Monjee Institute of Management Studies	136.8
2021	Sagi Ramakrishnam Raju Engineering College	118.3
2021	Shri Ramdeobaba College of Engineering and Management	110.8
2022	Alliance University	172.5
2022	Anna University	13.4
2022	B. S. Abdur Rahman Crescent Institute of Science and Technology	106.0
2022	B.M.S. College of Engineering	73.7
2022	Bansilal Ramnath Agarwal Charitable Trust's Vishwakarama Institute of Technology	155.0
2022	G. L. Bajaj Institute of Technology and Management	194.5

Continued on next page

Year	Institution	Avg. Rank
2022	Indian Institute of Information Technology, Design & Manufacturing, Kancheepuram	187.7
2022	Indian Institute of Technology Mandi	28.8
2022	Indian Institute of Technology Palakkad	66.3
2022	Jaypee University of Information Technology, Solan	111.7
2022	Lovely Professional University	68.1
2022	Maharaja Sayajirao University of Baroda	119.0
2022	Netaji Subhas University of Technology	71.7
2022	R.V. College of Engineering	75.1
2022	Sathyabama Institute of Science and Technology	55.4
2022	Saveetha Engineering College	107.5
2022	Sharda University	171.5
2022	Walchand College of Engineering	173.3
2023	Dr. B. R. Ambedkar National Institute of Technology, Jalandhar	62.4
2023	Galgotias University	162.0
2023	Indian Institute of Technology Bhubaneswar	31.0
2023	Institute of Infrastructure Technology Research and Management	190.0
2023	National Institute of Advanced Manufacturing Technology	137.7
2023	National Institute of Technology Delhi	88.8
2023	Shoolini University of Biotechnology and Management Sciences	108.0
2023	Silicon Institute of Technology, Bhubaneswar	167.7
2023	The LNM Institute of Information Technology	194.0
2024	Amity University Haryana, Gurgaon	106.5
2024	Amity University, Gwalior	156.3
2024	Amity University, Jaipur	133.0
2024	Army Institute of Technology	112.7
2024	Institute of Engineering & Management	153.7
2024	KIET Group of Institutions	181.5
2024	Mahindra University	154.0
2024	Manav Rachna International Institute of Research & Studies	140.3
2024	NMAM Institute of Technology	147.0
2024	Sona College of Technology	154.3
2024	Thiagarajar College of Engineering	58.7
	Anurag University	158.8
	Hindustan Institute of Technology and Science	131.5
	Jaypee Institute of Information Technology	81.3
	National Institute of Technology Warangal	24.2
	PES University	99.7
	Padmasri Dr. B.V. Raju Institute of Technology	168.0
	Shri Mata Vaishno Devi University	87.4

Note: This table lists the 87 institutions that established an AI center between 2016 and 2024, including 7 institutions for which the year of establishment is unknown.

Institutions With Unknown Years As shown in Table B8, there are seven institutions without information on establishment year. To assess whether institutions with missing establishment years differ systematically from those for which we were able to verify the year of AI center establishment, Table B9 compares characteristics across the two groups. For four-year undergraduate programs, the two sets of institutions are similar in overall size and in the shares of female and within-state students. The only statistically signifi-

cant difference is a lower share of socially challenged students among institutions with unknown years. For two-year master's programs, the groups also appear broadly comparable, although institutions with confirmed years have a smaller share of female students. Differences in capital and operating expenditures are generally small and imprecisely estimated, with the exception of slightly higher library and equipment spending among institutions with confirmed establishment years. To assess overall balance across these observables, we conduct a joint F-test of equality of means, which yields a p-value of 0.18, indicating that we cannot reject the null hypothesis that the two groups are similar across these characteristics. Taken together, the results suggest that institutions with missing establishment years do not differ meaningfully from those with confirmed dates, and the absence of establishment-year information is unlikely to introduce systematic bias into our analysis.

TABLE B9: Summary Statistics: Confirmed vs. Unknown Establishment Year (2016-2018, Institutions with AI Centers)

	(1)	(2)	(3)	(4)	(5)
	Confirmed		Unknown		t-test
Variable	N	Mean/(SD)	N	Mean/(SD)	Difference
Four-Year Undergraduate Programs					
Total No. of students	119	3234.597	14	3685.071	-450.475
		(2449.820)		(1377.007)	
Share of female	119	0.216	14	0.252	-0.035
		(0.138)		(0.050)	
Share of within-state	119	0.503	14	0.568	-0.064
		(0.344)		(0.210)	
Share of socially challenged	119	0.415	14	0.245	0.170***
		(0.202)		(0.182)	
Two-Year Master's Programs					
Total No. of students	120	634.917	14	388.000	246.917
		(661.593)		(373.206)	
Share of female	120	0.320	14	0.430	-0.110**
		(0.155)		(0.185)	
Share of within-state	120	0.543	14	0.699	-0.157
		(0.358)		(0.221)	
Share of socially challenged	120	0.373	14	0.281	0.092
		(0.230)		(0.172)	
Capital Expenditures (million INR)					
Library	185	25.163	21	7.932	17.231*
		(42.491)		(8.912)	
Equipment	185	103.675	21	32.288	71.387
		(234.700)		(43.176)	
Eng. Workshops	185	17.088	21	6.573	10.515
		(44.910)		(9.958)	
Other Cap.	185	53.889	21	21.055	32.835
		(120.737)		(34.727)	
Operating Expenditures (million INR)					
Salaries	185	495.788	21	324.321	171.468
		(643.038)		(202.436)	
Maintenance	185	319.882	21	225.036	94.846
		(500.295)		(208.681)	
Seminars	185	11.225	21	17.252	-6.026
		(29.340)		(27.549)	

Notes: We restrict the sample to institutions with AI centers and compare institutions with a confirmed establishment year versus those without a confirmed establishment year. "Library" = The annual capital expenditure on library including books, journals, and e-Resources. "Equipment" = The annual capital expenditure on new equipment and software for laboratories. "Workshops" = The annual capital expenditure on engineering workshops. "Salaries" = The annual operating expenditure on salaries for teaching and non-teaching staff. "Maintenance" = The annual operating expenditure on the maintenance of academic infrastructure, consumables, and other running expenses (excluding maintenance of hostels and allied services, building rent, depreciation costs, etc.). "Seminar" = The annual operating expenditure on seminars, conferences, and workshops. All expenditure variables are measured in million INR. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. We also conduct a joint F-test of equality of means. The p-value is 0.18.

C Data Construction

In this section, we provide details on the construction of our dataset and the definitions of all variables used in the analysis.

NIRF Data Our primary source of data is the NIRF, which annually publishes rankings and detailed performance metrics for higher education institutions in India, typically in September of each year. Established by the Ministry of Education in 2015, the NIRF provides a standardized framework for evaluating and comparing institutions across disciplines such as engineering, management, pharmacy, and the arts and sciences. Our analysis focuses on the engineering discipline rankings. It is important to note that not all institutions participating in the NIRF receive a discipline-specific ranking. Institutions appear in the engineering rankings only if they offer full-time engineering programs and have graduated at least three batches of students in those programs.

Ranking Process The ranking process is based on five broad parameters, each composed of several quantitative indicators derived from institution-reported data and publicly verifiable sources: (1) Teaching, Learning, and Resources (TLR), which includes indicators such as faculty-student ratio and faculty qualifications; (2) Research and Professional Practice (RP), which covers publications and patents; (3) Graduation Outcome (GO), which measures student placement rates and median salaries; (4) Outreach and Inclusivity (OI), which reflects gender and social diversity; and (5) Perception (PR), based on surveys of employers, academics, and the general public.

Collation of Data We use the annual data reports published under the Engineering discipline rankings on the official NIRF website between 2017 and 2025, covering 268 unique engineering institutions. We exclude the 2016 data because the reported statistics do not separately identify graduation outcomes for four-year undergraduate and two-year postgraduate programs. All statistics are reported at the institution level. Each year's ranking is based on the previous academic year; for example, the 2024 ranking uses data from the 2022-23 academic year. In our dataset, the 2022-23 academic year is coded as 2023, reflecting the year in which placement outcomes are realized; the same convention is applied to all other years.

For the rankings published between 2019 and 2022, NIRF released detailed institutional data for the top 200 engineering colleges, whereas for 2017-2018 and 2023-2025, detailed reports are available only for the top 100 institutions. Institutions that remain

consistently ranked within the top 100 therefore have complete data for all years in our sample, yielding 56 institutions. However, many institutions move above and below the top-100 cutoff over time. When such movements occur in 2017, 2018, 2023, 2024, or 2025—years in which NIRF publishes detailed data only for the top 100 institutions—the corresponding data reports for those institutions are not released, creating gaps in the panel.

To address this missingness, we take advantage of a key feature of NIRF submissions. Beginning with the 2019 ranking cycle, institutions are required to report three academic years of information on placement outcomes, capital expenditure, and operating expenditure. For sponsored research and consultancy projects, data collection begins in the 2021 ranking cycle, with institutions again required to report the previous three academic years. As a result, each institution's PDF submission in 2019-2025 includes up to two earlier years of data, even if the institution does not appear in the top-100 or top-200 list for that particular year. By extracting these retrospective fields from each PDF, we can recover missing observations and reconstruct a more complete panel.[36] This procedure increases the number of institutions with complete nine-year data to 78, of which 75 have an average engineering rank of 100 or better.

Summary Statistics Table C10 provides definitions of all variables used in the analysis. Table C11 presents summary statistics comparing treat institutions, those that establish an AI center between 2017 and 2024, with control institutions that do not, focusing only on the elite institutions in our sample. As shown, treated institutions enroll fewer undergraduate students but have higher postgraduate enrollment. They also exhibit higher levels of capital and operating expenditures, particularly on laboratory equipment and software, and salaries for teaching and non-teaching staff. In terms of student composition, treated institutions enroll a slightly higher share of socially disadvantaged students, and a lower share of within-state and female students.

[36]To validate the retrospective backfilled fields, we randomly selected several institutions from the 56 that already have complete data without backfilling and compared the overlapping years reported across multiple PDF submission cycles. The values match exactly in most cases and, where they differ, the discrepancies are small and immaterial. This consistency reassures us that the backfilled fields reflect the same underlying administrative records rather than ex post revisions.

TABLE C10: Variable Definition

Variable	Definition
Four-year undergraduate programs	
% Higher Studies	The ratio of graduates from four-year undergraduate programs pursuing higher studies to the total number of students graduating
% Placed	The placement rate = The ratio of students from four-year undergraduate programs who are placed to the total number of students graduating in that year
Salary (1,000 INR)	The median annual salary of employed graduates from four-year undergraduate programs, expressed in thousands of INR
Two-year master's programs	
% Higher Studies	The ratio of graduates from two-year master's programs pursuing higher studies to the total number of students graduating
% Placed	The placement rate = The ratio of students from two-year master's programs who are placed to the total number of students graduating in that year
Salary (1,000 INR)	The median annual salary of employed graduates from two-year master's programs, expressed in thousands of INR
Capital expenditure (million INR)	
Library	Annual capital expenditure on library including books, journals, and e-resources
Equipment	Annual capital expenditure on new equipment and software for laboratories
Eng. Workshops	Annual capital expenditure on establishing or upgrading engineering workshops, including machinery and tools
Other Cap.	Annual capital expenditure on the creation of other capital assets, such as setting up classrooms, seminar halls, conference halls, libraries, laboratories, and engineering workshops, excluding expenditure on land and buildings
Operating expenditure (million INR)	
Salaries	Annual operating expenditure on salaries for teaching and non-teaching staff
Maintenance	Annual operating expenditure on the maintenance of academic infrastructure, consumables, and other running expenses (excluding maintenance of hostels and allied services, building rent, depreciation costs, etc.)
Seminars	Annual operating expenditure on organizing seminars, conferences, and workshops, including logistics and materials
Sponsored research projects	
# of Projects	The total number of sponsored research projects undertaken
# of Agencies	The total number of funding agencies

Continued on next page

Variable	Definition
Total Amount (million INR)	The total amount of money received for the sponsored research projects, expressed in millions of INR
Consultancy projects	
# of Projects	The total number of consultancy projects undertaken
# of Agencies	The total number of client organizations
Total Amount (million INR)	The total amount of money received for the consultancy projects, expressed in millions of INR
Student composition (four-year undergraduate programs)	
Share of within-state	The ratio of within-state students enrolled across all years of four-year undergraduate programs to the total number of students enrolled in four-year undergraduate programs
Share of socially challenged	The ratio of socially challenged (Scheduled Caste) students enrolled across all years of four-year undergraduate programs to the total number of students enrolled in four-year undergraduate programs
Share of female	The ratio of female students enrolled across all years of four-year undergraduate programs to the total number of students enrolled in four-year undergraduate programs

TABLE C11: Summary Statistics: Control vs. Treated Institutions (2016-2018)

	(1)	(2) Control	(3)	(4) Treat	(5) t-test
Variable	N	Mean/(SD)	N	Mean/(SD)	Difference
Four-Year Undergraduate Programs					
Total No. of students	170	4185.576	77	3486.610	698.966
		(3967.127)		(2798.364)	
Share of female	170	0.221	77	0.184	0.038**
		(0.099)		(0.126)	
Share of within-state	170	0.515	77	0.403	0.112***
		(0.287)		(0.325)	
Share of socially challenged	170	0.383	77	0.440	-0.057**
		(0.207)		(0.183)	
Two-Year Master's Programs					
Total No. of students	175	572.669	79	844.329	-271.661***
		(514.804)		(718.272)	
Share of female	175	0.333	79	0.285	0.047**
		(0.139)		(0.136)	
Share of within-state	175	0.518	79	0.407	0.112**
		(0.315)		(0.339)	
Share of socially challenged	175	0.383	79	0.413	-0.030
		(0.201)		(0.216)	
Capital Expenditures (million INR)					
Library	211	18.298	87	47.969	-29.671***
		(20.582)		(52.705)	
Equipment	211	77.496	87	202.506	-125.010***
		(90.396)		(314.565)	
Eng. Workshops	211	11.610	87	30.518	-18.908***
		(28.997)		(62.054)	
Other Cap.	211	87.478	87	87.556	-0.078
		(208.214)		(157.546)	
Operating Expenditures (million INR)					
Salaries	211	481.937	87	823.269	-341.332***
		(444.247)		(814.517)	
Maintenance	211	346.411	87	558.743	-212.332***
		(496.843)		(644.739)	
Seminars	211	14.989	87	19.027	-4.038
		(41.208)		(41.142)	

Notes: We restrict the sample to institutions with an average NIRF engineering rank of 100 or better, resulting in 113 academic institutions. Treated institutions are defined as those that established an AI center between 2017 and 2024, whereas control institutions are those that did not establish an AI center over the same period. "Library" = The annual capital expenditure on library including books, journals, and e-Resources. "Equipment" = The annual capital expenditure on new equipment and software for laboratories. "Workshops" = The annual capital expenditure on engineering workshops. "Salaries" = The annual operating expenditure on salaries for teaching and non-teaching staff. "Maintenance" = The annual operating expenditure on the maintenance of academic infrastructure, consumables, and other running expenses (excluding maintenance of hostels and allied services, building rent, depreciation costs, etc.). "Seminar" = The annual operating expenditure on seminars, conferences, and workshops. All expenditure variables are measured in million INR. All monetary variables are converted to real terms using the CPI (2016 = 100). Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D More Details of the SDID Estimation Strategy

This appendix provides additional detail on the implementation of Synthetic Difference-in-Differences (SDID).

D.1 Staggered Adoption Design

We implement SDID with staggered treatment adoption by estimating cohort-specific SDID models. For each adoption year $a \in A$, where A denotes the set of unique adoption years, we follow the procedure outlined in Clarke et al. (2024):

1. For a given adoption year a, we restrict the sample to institutions that establish an AI center in year a and to institutions that have not yet established one by year a, including never-adopting institutions. The latter group serves as the donor pool. We denote this subsample by $\delta^{(a)}$ and refer to it as the analysis sample for adoption cohort a.

2. For each cohort a, we compute the cohort-specific regularization parameter $\zeta^{(a)}$, the cohort-specific unit weights $\hat{\omega}^{sdid}_{(a)}$, and the cohort-specific time weights $\hat{\lambda}^{sdid}_{(a)}$ laid out in Arkhangelsky et al. (2021).

3. For each adoption cohort a, we estimate the SDID treatment effect by solving the following regularized weighted least-squares problem:

$$(\hat{\tau}^{sdid}_{(a)}, \hat{\mu}_{(a)}, \hat{\alpha}_{(a)}, \hat{\beta}_{(a)}) = \arg\min_{\tau,\mu,\alpha,\beta} \left\{ \sum_{i \in \delta^{(a)}} \sum_{t=1}^{T} (Y_{it} - \mu - \alpha_i - \beta_t - \tau D_{i,t})\, \hat{\omega}^{sdid}_{(a),i} \hat{\lambda}^{sdid}_{(a),t} \right\},$$

   where Y_{it} denotes the outcome for institution i in year t, μ is an intercept, α_i are institution fixed effects, γ_t are year fixed effects, and $D_{i,t}$ is the treatment indicator for institution i in year t. The parameter $\hat{\tau}^{sdid}_{(a)}$ represents the estimated average treatment effect on treated for adoption cohort a.

4. The overall average treatment effect on the treated is computed as a weighted average of the cohort-specific effects:

$$\widehat{ATT} \text{ or } \hat{\tau}^{sdid} = \sum_{a \in A} \frac{T^{(a)}_{post}}{T_{post}} \times \hat{\tau}^{sdid}_{(a)},$$

   where $T^{(a)}_{post}$ denotes the number of post-treatment periods observed for adoption

cohort a, and T_{post} denotes the total number of post-treatment periods observed among treated institutions.

D.2 Additional Standard Errors for SDID

We also report placebo-inference standard errors (shown in square brackets) in Tables 2-6, based on 999 replications in which treatment is randomly reassigned to never-treated institutions within each cohort's adoption year. This serves as an additional robustness check. Although placebo inference is often recommended when the number of treated units is small, it assumes homoskedasticity across units and is therefore not our preferred approach. We do not use jackknife inference because, as shown by the yellow bars in Figure 2, there is only one treated institution in each of the years 2017, 2018, and 2024. Since the jackknife removes one unit at a time, these years would have no remaining treated observations, making the estimator undefined.

D.3 SDID Weights

The SDID estimator combines unit weights ($\hat{\omega}_i$) and time weights ($\hat{\lambda}_t$) to construct synthetic controls that closely reproduce the pre-treatment path of each treatment cohort. Because treatment occurs in different years across institutions, SDID solves a separate optimization problem for each first-treatment year, generating cohort-specific weights. As an illustration, we report the unit and time weights for the three graduate outcome variables for four-year undergraduate programs.

Table D12 presents the distribution of unit weights that SDID assigns to never-treated institutions for each adoption cohort. As shown, not all never-treated institutions receive positive weight. The institutions that do receive positive weights are those whose pre-treatment trajectories most closely resemble those of the treated cohort. The average magnitude of these positive weights across adoption cohorts and outcome variables is between 0.02 and 0.03.

Table D13 reports the distribution of time weights assigned to the pre-treatment years for each cohort. These weights also vary across adoption years and outcome variables. Pre-treatment years that more closely align with a cohort's baseline trajectory receive larger weights. In the post-treatment period, SDID imposes equal time weights, with each post-treatment year receiving a weight of 1/the number of post-treatment years.

Together, the unit and time weights illustrate how SDID constructs synthetic comparison groups tailored to each cohort. The unit weights reweigh the most comparable never-treated institutions, while the time weights emphasize pre-treatment periods that

best fit the observed baseline trajectory. These cohort-specific weights provide evidence of a close pre-treatment match, supporting the credibility of the SDID counterfactuals used in the main analysis.

TABLE D12: SDID Unit Weights: Four-Year Undergraduate

Adoption Year	(1) No. of Institutions	(2) Mean	(3) SD	(4) Min	(5) Max
Panel A. Dependent variable = % Higher Studies					
2017	48	0.021	0.000	0.021	0.021
2018	40	0.025	0.019	0.001	0.073
2019	48	0.021	0.005	0.005	0.032
2020	45	0.022	0.008	0.006	0.041
2021	45	0.022	0.017	0.001	0.086
2022	45	0.022	0.010	0.002	0.044
2023	37	0.027	0.021	0.000	0.084
Panel B. Dependent variable = % Placed					
2017	48	0.021	0.000	0.021	0.021
2018	48	0.021	0.005	0.007	0.030
2019	48	0.021	0.005	0.007	0.033
2020	48	0.021	0.007	0.004	0.036
2021	43	0.023	0.008	0.010	0.043
2022	42	0.024	0.016	0.003	0.073
2023	38	0.026	0.016	0.005	0.066
Panel C. Dependent variable = Salary (1,000 INR)					
2017	48	0.021	0.000	0.021	0.021
2018	42	0.024	0.022	0.001	0.096
2019	48	0.021	0.005	0.005	0.030
2020	33	0.030	0.026	0.002	0.109
2021	42	0.024	0.020	0.000	0.071
2022	35	0.029	0.022	0.001	0.078
2023	31	0.032	0.025	0.000	0.080

Notes: In column (1), we report the number of never-treated institutions that receive a positive unit weight. Note that "0.000" reflects a very small positive value rather than an exact zero.

TABLE D13: SDID Time Weights: Four-Year Undergraduate

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Pre-Treatment Year						
Adoption Year	2016	2017	2018	2019	2020	2021	2022
Panel A. Dependent variable = % Higher Studies							
2017	1.000						
2018	0.031	0.969					
2019	0.000	0.778	0.222				
2020	0.021	0.637	0.086	0.256			
2021	0.067	0.605	0.008	0.138	0.182		
2022	0.140	0.535	0.006	0.032	0.040	0.247	
2023	0.000	0.035	0.000	0.285	0.000	0.000	0.680
Panel B. Dependent variable = % Placed							
2017	1.000						
2018	0.473	0.527					
2019	0.330	0.058	0.612				
2020	0.268	0.018	0.332	0.382			
2021	0.225	0.044	0.284	0.121	0.326		
2022	0.041	0.000	0.290	0.062	0.000	0.607	
2023	0.149	0.000	0.079	0.000	0.000	0.400	0.373
Panel C. Dependent variable = Salary (1,000 INR)							
2017	1.000						
2018	0.343	0.657					
2019	0.096	0.207	0.697				
2020	0.030	0.000	0.074	0.896			
2021	0.000	0.028	0.048	0.350	0.575		
2022	0.000	0.000	0.000	0.178	0.563	0.259	
2023	0.143	0.000	0.000	0.000	0.299	0.186	0.372

Notes: We report the pre-treatment time weights separately for each adoption cohort. In the post-treatment period, the time weights are uniform and equal to $\frac{1}{\#\text{ of post-treatment years}}$. Values reported as "0.000" are not necessarily exact zeros; they may reflect very small positive weights rounded to three decimals.

E Additional Results

TABLE E1: Predicted Adoption Risk and Pre-Treatment Trends

	(1) % Placed Pre-Trend	(2) Salary (1,000 INR) Pre-Trend
Predicted Adoption Risk	-0.001	-1.912
	(0.004)	(3.150)
R^2	0.01	0.01
Observations	73	73

Notes: The table tests whether predicted AI-center adoption is associated with differential pre-treatment trends. Predicted adoption risk is constructed from a discrete-time hazard model of AI-center adoption, estimated on the 2016–2024 institution-year panel using a complementary log-log specification with year and state fixed effects. The sample is restricted to institutions with average NIRF rank 100 or better and excludes institutions treated in 2016 or earlier. Pre-treatment trends are estimated at the institution level using only clean pre-treatment years: for treated institutions, years at least two years before adoption; for never-treated institutions, years 2016-2018. Column 1 uses the institution-specific pre-trend in the placement share of four-year undergraduates; column 2 uses the institution-specific pre-trend in the median salary of placed four-year undergraduates. Standard errors are reported in parentheses. All monetary variables are converted to real terms using the CPI (2016 = 100). Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE E2: Effects of AI Center Establishment on Log Specifications of Graduate Outcomes: Four-Year Undergraduate Programs (2016-2024)

VARIABLES	(1)	(2)	(3)	(4)
	Ln(% Placed)		Ln(Salary (1,000 INR))	
	SDID	BJS	SDID	BJS
Treatment	0.072**	0.063**	0.038	0.042
	(0.032)	(0.025)	(0.034)	(0.031)
Placebo SE	[0.039]*		[0.042]	
2016 Mean	-0.33	-0.032	6.31	6.26
No. of Institutions	72	107	72	107
N	648	844	648	857

Notes: We employ three estimation methods on the institution-year panel data (2016-2024). Columns (1) and (3) report estimates from the synthetic difference-in-differences (SDID) method with staggered adoption (Arkhangelsky et al., 2021). Columns (2) and (4) report estimates from the DID imputation method (Borusyak et al., 2024). Standard errors, shown in parentheses, are obtained via bootstrap inference with 999 replications in columns (1) and (3), and are clustered at the institution level in columns (2) and (4). As a robustness check, we additionally report placebo-based standard errors with 999 replications for columns (1) and (3). "Ln(% Place)" = The logarithm of the placement rate, defined as the proportion of graduates from four-year undergraduate programs who obtained employment. "Ln(Salary (1,000 INR))" = The logarithm of the median salary of employed graduates from four-year undergraduate programs, expressed in thousands of Indian Rupees (1,000 INR). All monetary variables are converted to real terms using the CPI (2016 = 100). "2016 Mean" is the average value of the respective variable in 2016. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE E3: Effects of AI Center Establishment on Graduate Outcomes: Four-Year Undergraduate Programs (Callaway and Sant'Anna, 2021)

VARIABLES	(1) % Placed	(2) Salary (1,000 INR)
Treatment	0.041** (0.018)	31.488 (26.016)
2016 Mean	0.74	576.87
No. of Institutions	105	105
N	855	855

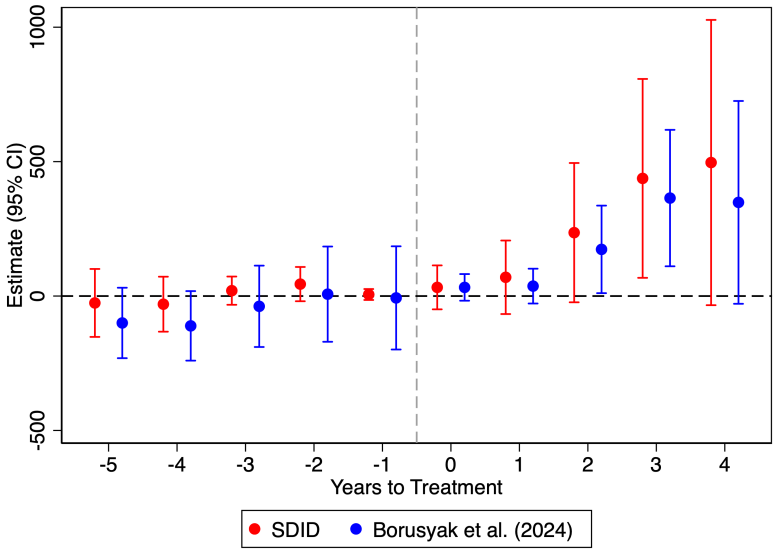
Notes: We employ the staggered DID method (Callaway and Sant'Anna, 2021) on the institution-year panel (2016–2024). Standard errors, shown in parentheses, are based on robust asymptotic inference. "% Place" = The proportion of graduates from four-year undergraduate programs who obtained employment. "Salary (1,000 INR)" = The median salary of employed graduates from four-year undergraduate programs, expressed in thousands of Indian Rupees (1,000 INR). All monetary variables are converted to real terms using the CPI (2016 = 100). "2016 Mean" is the average value of the respective variable in 2016. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE E4: Heterogeneous Effects of AI Center Establishment on Graduate Outcomes by **Share of Female Students**: Four-Year Undergraduate Programs

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	% Higher Studies		% Placed		Salary (1,000 INR)	
	SDID	BJS	SDID	BJS	SDID	BJS
Treatment (Above)	0.021	-0.001	0.041	0.022	17.332	-24.872
	(0.029)	(0.009)	(0.038)	(0.037)	(28.765)	(23.094)
Treatment (Below)	-0.025	-0.014	0.036	0.041**	53.938	96.396***
	(0.018)	(0.014)	(0.027)	(0.017)	(39.193)	(26.740)
Difference	0.046	0.014	0.005	-0.020	-36.607	-121.268***
	(0.034)	(0.014)	(0.047)	(0.039)	(48.616)	(27.948)
2016 Mean	0.11	0.10	0.73	0.74	604.90	576.87
No. of Institutions	72	94	72	94	72	94
N	648	777	648	777	648	777

Notes: We employ two estimation methods on the institution-year panel (2016–2024). Columns (1), (3), and (5) report estimates from the synthetic difference-in-differences (SDID) method with staggered adoption (Arkhangelsky et al., 2021). Columns (2), (4), and (6) report estimates from the DID imputation method (Borusyak et al., 2024). Standard errors, shown in parentheses, are obtained via bootstrap inference with 999 replications in columns (1), (3), and (5), and are clustered at the institution level in columns (2), (4), and (6). "Treatment (Above)" reports estimates for the subsample of institutions with an average share of female students above the median, while "Treatment (Below)" refers to estimates for those with an average share of female students below the median. "% Higher Studies" = The proportion of graduates from four-year undergraduate programs who pursue higher studies. "% Place" = The proportion of graduates from four-year undergraduate programs who obtained employment. "Salary (1,000 INR)" = The median salary of employed graduates from four-year undergraduate programs, expressed in thousands of Indian Rupees (1,000 INR). All monetary variables are converted to real terms using the CPI (2016 = 100). "2016 Mean" is the average value of the respective variable in 2016. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

FIGURE E1: Dynamic Effects on Salary Expenditures



Notes: We employ two estimation methods on the institution-year panel data (2016-2024). The figure reports event-study estimates of the impact of AI center establishment on salary expenditures. "Salaries" denotes annual operating expenditure on salaries for teaching and non-teaching staff, expressed in million INR. The monetary variable is converted to real terms using the CPI (2016 = 100). Red estimates are obtained using Synthetic Difference-in-Differences with staggered adoption (Arkhangelsky et al., 2021); and blue estimates are obtained using the Borusyak-Jaravel-Spiess imputation estimator (Borusyak et al., 2024). The horizontal black dashed line denotes zero. The point at Years to Treatment = 0 corresponds to the year in which an AI center was established. Confidence intervals are shown at the 95 percent level.

F Robustness Checks

F.1 Placebo Test

TABLE F1: Effects of AI Center Establishment on Four-Year Undergraduate Placement and Salary Outcomes (Placebo)

VARIABLES	(1) Estimate	(2) SE	(3) p-value	(4) p-value (Placebo)
% Placed	0.038	(0.016)	0.02	0.09
Salary (1,000 INR)	66.455	(21.975)	0.00	0.02

Notes: We employ the the DID imputation method (Borusyak et al., 2024) on the institution-year panel (2016-2024). Column (4) reports empirical p-values based on 1,000 placebo repetitions in which treatment years are randomly assigned to never-treated institutions. "% Place" = The proportion of graduates from four-year undergraduate programs who obtained employment. "Salary (1,000 INR)" = The median salary of employed graduates from four-year undergraduate programs, expressed in thousands of Indian Rupees (1,000 INR). All monetary variables are converted to real terms using the CPI (2016 = 100). Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

F.2 Full Sample: BJS Method

TABLE F2: Effects of AI Center Establishment on Graduate Outcomes: Four-Year Undergraduate Programs (Full Sample)

VARIABLES	(1) % Higher Studies	(2) % Placed	(3) Salary (1,000 INR)
Treatment	-0.010	0.032**	58.296**
	(0.009)	(0.016)	(22.905)
2016 Mean	0.11	0.70	458.88
No. of Institutions	253	253	253
N	1,560	1,557	1,560

Notes: We employ the DID imputation method (Borusyak et al., 2024) on the full institution-year panel (2016-2024). Standard errors, shown in parentheses, are clustered at the institution level. "% Higher Studies" = The proportion of graduates from four-year undergraduate programs who pursue higher studies. "% Place" = The proportion of graduates from four-year undergraduate programs who obtained employment. "Salary (1,000 INR)" = The median salary of employed graduates from four-year undergraduate programs, expressed in thousands of Indian Rupees (1,000 INR). All monetary variables are converted to real terms using the CPI (2016 = 100). "2016 Mean" is the average value of the respective variable in 2016. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE F3: Impact of AI Center Establishment on Capital & Operating Expenditures (Full Sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Capital Expenditures				Operating Expenditures		
VARIABLES	Library	Equipment	Eng. Workshops	Other Cap.	Salaries	Maintenance	Seminars
Treatment	-1.528	74.681**	-12.530***	13.482	165.604***	141.646***	0.430
	(1.620)	(30.234)	(4.785)	(16.982)	(48.935)	(52.646)	(3.034)
2016 Mean	13.86	60.49	10.82	42.01	363.96	230.10	7.37
No. of Institutions	256	256	256	256	256	256	256
N	1,601	1,601	1,601	1,601	1,601	1,601	1,601

Notes: We employ the DID imputation method (Borusyak et al., 2024) on the full institution-year panel (2016-2024). Standard errors, shown in parentheses, are clustered at the institution level. "Library" = The annual capital expenditure on library including books, journals, and e-Resources. "Equipment" = The annual capital expenditure on new equipment and software for laboratories. "Eng. Workshops" = The annual capital expenditure on establishing or upgrading engineering workshops, including machinery and tools. "Other Cap." = The annual capital expenditure on the creation of other capital assets, such as setting up classrooms, seminar halls, conference halls, libraries, laboratories, and engineering workshops, excluding expenditure on land and buildings. "Salaries" = The annual operating expenditure on salaries for teaching and non-teaching staff. "Maintenance" = The annual operating expenditure on the maintenance of academic infrastructure, consumables, and other running expenses (excluding maintenance of hostels and allied services, building rent, depreciation costs, etc.). "Seminar" = The annual operating expenditure on organizing seminars, conferences, and workshops, including event logistics, materials, and related recurring costs. The units of all outcome variables are in million INR. All monetary variables are converted to real terms using the CPI (2016 = 100). "2016 Mean" is the average value of the respective variable in 2016. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE F4: Impact of AI Center Establishment on Research Projects (Full Sample)

	(1)	(2)	(3)	(4)	(5)	(6)
	Sponsored Research			Consultancy Projects		
VARIABLES	# of Projects	# of Agencies	Total Amount (million INR)	# of Projects	# of Agencies	Total Amount (million INR)
Treatment	27.077	8.533	50.003	-27.315	-32.211	16.441
	(18.144)	(5.931)	(32.677)	(40.870)	(67.015)	(10.181)
2018 Mean	59.99	18.73	112.34	87.54	182.21	28.76
No. of Institutions	215	215	215	215	215	215
N	1,111	1,111	1,111	1,111	1,111	1,111

Notes: We employ the DID imputation method (Borusyak et al., 2024) on the full institution-year panel (2016-2024). Standard errors, shown in parentheses, are clustered at the institution level. Columns (1)-(3) report the results for sponsored research projects, while columns (4)-(6) present the results for consultancy projects. "# of Projects" = The total number of sponsored projects at the institution level in column (1) and the total number of consultancy projects in column (4) "# of Agencies" = The total number of funding agencies in column (2) and the total number of client organizations in column (5). "Total Amount (million INR)" = The total funding received for sponsored research projects in column (3) and the total amount received for consultancy projects in column (6). The units of outcome variables in columns 3 and 6 are in million INR. All monetary variables are converted to real terms using the CPI (2016 = 100). "2018 Mean" is the average value of the respective variable in 2018. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

F.3 Restricting to Top 50 Institutions

TABLE F5: Effects of AI Center Establishment on Student Outcomes: Four-Year Undergraduate Programs (Top-50 Sample)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	% Higher Studies		% Placed		Salary (1,000 INR)	
	SDID	BJS	SDID	BJS	SDID	BJS
Treatment	-0.006	-0.003	0.057**	0.055***	23.271	44.306
	(0.020)	(0.011)	(0.027)	(0.019)	(35.392)	(29.106)
Placebo SE	[0.022]		[0.032]*		[34.248]	
2016 Mean	0.11	0.10	0.74	0.74	689.42	689.42
No. of Institutions	45	107	45	46	45	46
N	405	413	405	413	405	413

Notes: We employ two estimation methods on the institution-year panel (2016-2024), restricting the sample to institutions with an average ranking of 50 or better. Columns (1), (3), and (5) report estimates from the synthetic difference-in-differences (SDID) method with staggered adoption (Arkhangelsky et al., 2021). Columns (2), (4), and (6) report estimates from the DID imputation method (Borusyak et al., 2024). Standard errors, shown in parentheses, are obtained via bootstrap inference with 999 replications in columns (1), (3), and (5), and are clustered at the institution level in columns (2), (4), and (6). As a robustness check, we additionally report placebo-based standard errors with 999 replications for columns (1), (3), and (5). "% Higher Studies" = The proportion of graduates from four-year undergraduate programs who pursue higher studies. "% Place" = The proportion of graduates from four-year undergraduate programs who obtained employment. "Salary (1,000 INR)" = The median salary of employed graduates from four-year undergraduate programs, expressed in thousands of Indian Rupees (1,000 INR). All monetary variables are converted to real terms using the CPI (2016 = 100). "2016 Mean" is the average value of the respective variable in 2016. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE F6: Impact of AI Center Establishment on Capital & Operating Expenditures (Top-50 Sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Capital Expenditures			Operating Expenditures		
VARIABLES	Library	Equipment	Eng. Workshops	Other Cap.	Salaries	Maintenance	Seminars
Panel A. SDID							
Treatment	2.456	189.601*	-28.484	1.398	360.987**	157.868	-4.320
Bootstrapped SE	(4.777)	(115.224)	(28.959)	(36.102)	(169.710)	(202.598)	(13.496)
Placebo SE	[4.665]	[40.807]***	[10.876]***	[46.038]	[117.763]***	[144.617]	[8.942]
2016 Mean	39.34	188.64	35.01	93.12	815.27	558.64	23.41
No. of Institutions	45	45	45	45	45	45	45
N	405	405	405	405	405	405	405
Panel B. BJS							
Treatment	0.108	146.851***	-35.837***	7.278	367.693***	294.682***	2.789
	(3.664)	(52.615)	(7.244)	(30.386)	(97.979)	(105.870)	(9.154)
2016 Mean	38.60	184.55	34.26	91.55	800.05	546.84	22.96
No. of Institutions	47	47	47	47	47	47	47
N	416	416	416	416	416	416	416

Notes: We employ two estimation methods on the institution-year panel (2016–2024), restricting the sample to institutions with an average ranking of 50 or better. Panel A reports estimates from the synthetic difference-in-differences (SDID) method with staggered adoption (Arkhangelsky et al., 2021). Panel B reports estimates from the DID imputation method (Borusyak et al., 2024). Standard errors, reported in parentheses, are obtained via bootstrap inference with 999 replications in Panel A, and are clustered at the institution level in Panel B. As a robustness check, Panel A also reports standard errors computed using placebo inference with 999 replications. "Library" = The annual capital expenditure on library including books, journals, and e-Resources. "Equipment" = The annual capital expenditure on new equipment and software for laboratories. "Eng. Workshops" = The annual capital expenditure on establishing or upgrading engineering workshops, including machinery and tools. "Other Cap." = The annual capital expenditure on the creation of other capital assets, such as setting up classrooms, seminar halls, conference halls, libraries, laboratories, and engineering workshops, excluding expenditure on land and buildings. "Salaries" = The annual operating expenditure on salaries for teaching and non-teaching staff. "Maintenance" = The annual operating expenditure on the maintenance of academic infrastructure, consumables, and other running expenses (excluding maintenance of hostels and allied services, building rent, depreciation costs, etc.). "Seminar" = The annual operating expenditure on organizing seminars, conferences, and workshops, including event logistics, materials, and related recurring costs. The units of all outcome variables are in million INR. All monetary variables are converted to real terms using the CPI (2016 = 100). "2016 Mean" is the average value of the respective variable in 2016. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE F7: Impact of AI Center Establishment on Research Projects (Top-50 Sample)

	(1)	(2)	(3)	(4)	(5)	(6)
	Sponsored Research			Consultancy Projects		
VARIABLES	# of Projects	# of Agencies	Total Amount (million INR)	# of Projects	# of Agencies	Total Amount (million INR)
Panel A. SDID						
Treatment	70.964	26.235*	113.653	0.186	73.319	55.187
Bootstrapped SE	(46.542)	(15.150)	(83.717)	(62.667)	(45.992)	(46.786)
Placebo SE	[30.009]***	[6.093]***	[33.890]***	[95.949]	[102.748]	[17.487]***
2018 Mean	147.66	35.77	360.91	129.95	252.82	75.84
No. of Institutions	44	44	44	44	44	44
N	308	308	308	308	308	308
Panel B. BJS						
Treatment	57.885	23.932*	7.619	-6.524	60.735	68.472
	(42.717)	(12.526)	(93.007)	(52.586)	(48.978)	(44.009)
2018 Mean	167.46	40.52	451.98	137.96	264.91	100.11
No. of Institutions	44	44	44	106	44	44
N	308	308	308	308	308	308

Notes: We employ two estimation methods on the institution-year panel (2018–2024), restricting the sample to institutions with an average ranking of 50 or better. Panel A reports estimates from the synthetic difference-in-differences (SDID) method with staggered adoption (Arkhangelsky et al., 2021). Panel B reports estimates from the DID imputation method (Borusyak et al., 2024). Standard errors, reported in parentheses, are obtained via bootstrap inference with 999 replications in Panel A, and are clustered at the institution level in Panel B. As a robustness check, Panel A also reports standard errors computed using placebo inference with 999 replications. Columns (1)-(3) report the results for sponsored research projects, while columns (4)-(6) present the results for consultancy projects. "# of Projects" = The total number of sponsored projects at the institution level in column (1) and the total number of consultancy projects in column (4) "# of Agencies" = The total number of funding agencies in column (2) and the total number of client organizations in column (5). "Total Amount (million INR)" = The total funding received for sponsored research projects in column (3) and the total amount received for consultancy projects in column (6). The units of outcome variables in columns 3 and 6 are in million INR. All monetary variables are converted to real terms using the CPI (2016 = 100). "2018 Mean" is the average value of the respective variable in 2018. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

F.4 Student Compositions: Share of Within-State Students

TABLE F8: Effects of AI Center Establishment on Share of Within-State Students: Four-Year Undergraduate Programs

VARIABLES	(1)	(2)
	Share of Within-State Students	
	SDID	BJS
Treatment	0.016	0.031
	(0.019)	(0.019)
Placebo SE	[0.022]	
2016 Mean	0.43	0.50
No. of Institutions	56	107
N	504	794

Notes: We employ two estimation methods on the institution-year panel. Column (1) reports estimates from the Synthetic Difference-in-Differences (SDID) method with staggered adoption (Arkhangelsky et al., 2021) and column (2) reports estimates from the DID imputation method (Borusyak et al., 2024). Standard errors, reported in parentheses, are obtained via bootstrap inference with 999 replications in column (3), and are clustered at the institution level in column (2) For column (1), we additionally report placebo-inference standard errors based on 999 replications as a robustness check. "2016 Mean" is the average value of the respective variable in 2016. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

F.5 Additional Controls: BJS Method

TABLE F9: Effects of AI Center Establishment on Graduate Outcomes: Four-Year Undergraduate Programs (Additional Controls)

VARIABLES	(1) % Higher Studies	(2) % Placed	(3) Salary (1,000 INR)
Panel A. Including state fixed effects			
Treatment	-0.010	0.038**	66.455***
	(0.011)	(0.016)	(21.975)
2016 Mean	0.10	0.74	576.87
State FE	Yes	Yes	Yes
State-Year FE	No	No	No
No. of Institutions	107	107	107
N	857	857	857
Panel B. Including district-level characteristics × year dummies			
Treatment	-0.013	0.034**	61.785***
	(0.012)	(0.015)	(21.333)
2016 Mean	0.10	0.74	576.87
State FE	No	No	No
State-Year FE	Yes	Yes	Yes
No. of Institutions	107	107	107
N	857	857	857

Notes: We employ the DID imputation method (Borusyak et al., 2024) on the institution-year panel dataset (2016-2024). Standard errors, shown in parentheses, are clustered at the institution level. Panel A additionally includes state fixed effects, whereas Panel B incorporates district-level labor-market characteristics from 2017 interacted with year dummies. These characteristics, drawn from the Periodic Labor Force Survey, include the work-population ratio (the number of employed persons relative to the population), the labor-force participation rate (the number of people in the labor force relative to the population), and the unemployment rate (the number of unemployed persons relative to the labor force). "% Higher Studies" = The proportion of graduates from four-year undergraduate programs who pursue higher studies. "% Place" = The proportion of graduates from four-year undergraduate programs who obtained employment. "Salary (1,000 INR)" = The median salary of employed graduates from four-year undergraduate programs, expressed in thousands of Indian Rupees (1,000 INR). All monetary variables are converted to real terms using the CPI (2016 = 100). "2016 Mean" is the average value of the respective variable in 2016. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE F10: Impact of AI Center Establishment on Capital & Operating Expenditures (Additional Controls)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Capital Expenditures				Operating Expenditures		
VARIABLES	Library	Equipment	Eng. Workshops	Other Cap.	Salaries	Maintenance	Seminars
Panel A. Including state fixed effects							
Treatment	-0.766	91.324***	-24.972***	16.805	216.798***	202.209***	0.954
	(2.216)	(33.969)	(5.018)	(23.684)	(60.786)	(65.199)	(4.948)
2016 Mean	23.84	109.98	18.78	78.23	551.41	378.77	12.95
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Year FE	No	No	No	No	No	No	No
No. of Institutions	110	110	110	110	110	110	110
N	893	893	893	8931	893	893	893
Panel B. Including district-level characteristics × year dummies							
Treatment	-0.765	92.488***	-24.442***	26.633	207.994***	215.864***	-0.481
	(2.410)	(34.792)	(5.047)	(24.835)	(61.704)	(67.287)	(4.854)
2016 Mean	23.84	109.98	18.78	78.23	551.41	378.77	12.95
State FE	No	No	No	No	No	No	No
State-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of Institutions	110	110	110	110	110	110	110
N	893	893	893	893	893	893	8937

Notes: We employ the DID imputation method (Borusyak et al., 2024) on the institution-year panel dataset (2016-2024). Standard errors, shown in parentheses, are clustered at the institution level. Panel A additionally include state fixed effects, while Panel B incorporates district-level labor-market characteristics from 2017 interacted with year dummies. These characteristics, drawn from the Periodic Labor Force Survey, include the work-population ratio (the number of employed persons relative to the population), the labor-force participation rate (the number of people in the labor force relative to the population), and the unemployment rate (the number of unemployed persons relative to the labor force). "Library" = The annual capital expenditure on library including books, journals, and e-Resources. "Equipment" = The annual capital expenditure on new equipment and software for laboratories. "Eng. Workshops" = The annual capital expenditure on establishing or upgrading engineering workshops, including machinery and tools. "Other Cap." = The annual capital expenditure on the creation of other capital assets, such as setting up classrooms, seminar halls, conference halls, libraries, laboratories, and engineering workshops, excluding expenditure on land and buildings. "Salaries" = The annual operating expenditure on salaries for teaching and non-teaching staff. "Maintenance" = The annual operating expenditure on the maintenance of academic infrastructure, consumables, and other running expenses (excluding maintenance of hostels and allied services, building rent, depreciation costs, etc.). "Seminar" = The annual operating expenditure on organizing seminars, conferences, and workshops, including event logistics, materials, and related recurring costs. The units of all outcome variables are in million INR. All monetary variables are converted to real terms using the CPI (2016 = 100). "2016 Mean" is the average value of the respective variable in 2016. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE F11: Impact of AI Center Establishment on Research Projects (Additional Controls)

	(1)	(2)	(3)	(4)	(5)	(6)
	Sponsored Research			Consultancy Projects		
VARIABLES	# of Projects	# of Agencies	Total Amount (million INR)	# of Projects	# of Agencies	Total Amount (million INR)
Panel A. Including state fixed effects						
Treatment	36.810	10.960	59.612	-28.395	-80.479	17.421*
	(25.600)	(8.192)	(39.390)	(53.357)	(99.852)	(10.374)
2018 Mean	90.50	26.32	204.78	134.70	233.97	49.92
No. of Institutions	106	106	106	106	106	106
N	682	682	682	682	682	682
Panel B. Including district-level characteristics × year dummies						
Treatment	38.151	12.531	66.698*	-20.911	-74.815	18.516*
	(25.762)	(8.390)	(39.767)	(54.551)	(97.137)	(10.877)
2018 Mean	90.50	26.32	204.78	134.70	233.97	49.92
No. of Institutions	106	106	106	106	106	106
N	682	682	682	682	682	682

Notes: We employ the DID imputation method (Borusyak et al., 2024) on the institution-year panel dataset (2016-2024). Standard errors, shown in parentheses, are clustered at the institution level. Panel A additionally include state fixed effects, while Panel B incorporates district-level labor-market characteristics from 2017 interacted with year dummies. These characteristics, drawn from the Periodic Labor Force Survey, include the work-population ratio (the number of employed persons relative to the population), the labor-force participation rate (the number of people in the labor force relative to the population), and the unemployment rate (the number of unemployed persons relative to the labor force). Columns (1)-(3) report the results for sponsored research projects, while columns (4)-(6) present the results for consultancy projects. "# of Projects" = The total number of sponsored projects at the institution level in column (1) and the total number of consultancy projects in column (4) "# of Agencies" = The total number of funding agencies in column (2) and the total number of client organizations in column (5). "Total Amount (million INR)" = The total funding received for sponsored research projects in column (3) and the total amount received for consultancy projects in column (6). The units of outcome variables in columns 3 and 6 are in million INR. All monetary variables are converted to real terms using the CPI (2016 = 100). "2018 Mean" is the average value of the respective variable in 2018. Significance levels are indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

G Stylized Conceptual Model

This section formalizes the two-barrier conceptual framework described in Section 5.2. The model captures how institutional differences in baseline preparedness, outside-option strength, and market access generate the heterogeneous treatment patterns estimated in the main text.

Institution Characteristics Let institution i differs along three pre-treatment characteristics:

- θ_i: baseline AI-complementary preparedness capturing institutional selectivity and academic strength;
- ψ_i: outside-option index capturing access to high-return labor markets (higher ψ_i means better outside options); and
- m_i: geographic mobility index capturing cohort's ability to high-return AI labor markets (higher m_i means greater mobility or lower access frictions).

Skill Production AI-center establishment introduces a productivity shock $A > 0$ that increases the effectiveness of AI-skill formation. Let the institution-level stock of employable AI skills be

$$S_i(A) = f(\theta_i, A), \qquad f_\theta > 0,\ f_A > 0,\ f_{\theta A} > 0.$$

The complementarity $f_{\theta A} > 0$ captures skill-biased technological change: better-prepared institutions convert the same AI-center investment into larger skill gains.

Barrier 1: Placement Within institution i, graduates face idiosyncratic placement frictions or match noise ε_{ij} (e.g., interview performance, recruiter access). Assume ε_{ij} are i.i.d. draws from a distribution with CDF $F(\cdot)$ and density $\phi(\cdot)$, independent of (θ_i, ψ_i). Graduate j obtains an AI-intensive job if

$$S_i + \varepsilon_{ij} \geq T,$$

where T is the threshold skill requirement for AI-intensive placement.

The institution-level placement rate is therefore

$$P_i(A) \equiv \Pr(S_i + \varepsilon_{ij} \geq T) = 1 - F\left(T - f(\theta_i, A)\right).$$

Differentiating the above equation with respect to A,

$$\frac{\partial P_i}{\partial A} = \phi\left(T - f(\theta_i, A)\right) \cdot f_A(\theta_i, A).$$

Because $f_{\theta A} > 0$, the marginal skill gain $f_A(\theta_i, A)$ is larger when θ_i is higher. Hence, holding the local density term fixed, institutions with higher baseline preparedness θ_i exhibit larger placement gains.

Barrier 2: Wage Returns In the absence of an AI center, graduates have an outside-option wage

$$w_i^{out} = w^{out}(\psi_i), \qquad \frac{\partial w^{out}}{\partial \psi_i} > 0,$$

so higher ψ_i corresponds to better counterfactual labor-market opportunities. For expositional simplicity, we begin by assuming that w_i^{out} depends only on ψ_i; in the next section, we show that our key comparative-static results continue to hold when the outside option also depends on market access m_i (i.e., $w_i^{out} = w^{out}(\psi_i, m_i)$).

Conditional on obtaining an AI-intensive job, wages incorporate the counterfactual outside option and an AI-specific return that depends on market access:

$$w_i^{AI}(A) = w^{out}(\psi_i) + g(S_i(A), m_i), \qquad g_S > 0,\ g_{Sm} > 0.$$

The complementarity $g_{Sm} > 0$ implies that the marginal return to AI skills is higher when mobility/market access is greater.

We define the proportional return to AI-sector entry relative to the outside option as

$$R_i(A) \equiv \frac{w_i^{AI}(A) - w^{out}(\psi_i)}{w^{out}(\psi_i)} = \frac{g(S_i(A), m_i)}{w^{out}(\psi_i)}.$$

Differentiating with respect to A,

$$\frac{\partial R_i}{\partial A} = \frac{1}{w^{out}(\psi_i)} \cdot \frac{\partial g(S_i(A), m_i)}{\partial A} = \frac{1}{w^{out}(\psi_i)} \cdot g_S\left(S_i(A), m_i\right) \cdot f_A(\theta_i, A).$$

Since $w^{out}(\psi_i)$ is increasing in ψ_i, the proportional wage gain from the AI-center shock is larger when outside options are weaker; equivalently, $\frac{\partial R_i}{\partial A}$ is decreasing in ψ_i. Meanwhile, because $g_{Sm} > 0$ implies $g_S(S_i(A), m_i)$ is increasing in m_i, $\frac{\partial R_i}{\partial A}$ is increasing in market access m_i. The first comparative static reflects the comparative-advantage implication

of Roy-type occupational choice models: conditional on selection into AI-intensive jobs, groups with weaker counterfactual opportunities experience larger wage gains from entry into a higher-return sector. The second comparative static captures that better access to high-return markets amplifies the payoff to AI skills conditional on placement.

Implications The model yields three testable predictions:

1. **Preparedness θ_i and placement**: Institutions with higher baseline preparedness (more selective, more mobile cohorts) exhibit larger post-treatment increases in placement rates; that is, $\frac{\partial P_i}{\partial A}$ is increasing in θ_i.

2. **Outside options ψ_i and conditional wage gains**: Conditional on AI-intensive placement, institutions with weaker outside options (historically disadvantaged cohorts) experience larger proportional wage gains; that is, $\frac{\partial R_i}{\partial A}$ is decreasing in ψ_i.

3. **Mobility m_i and conditional wage gains**: Conditional on AI-intensive placement, institutions with greater geographic mobility and market access experience larger proportional wage gains; that is, $\frac{\partial R_i}{\partial A}$ is increasing in m_i.

These predictions align with the empirical heterogeneity patterns documented in Section 5.3.

Alternative Form of Outside-Option Wage Suppose instead that, absent an AI center, graduates earn an outside-option wage

$$w_i^{out} = w^{out}(\psi_i, m_i), \qquad \frac{\partial w^{out}}{\partial \psi_i} > 0, \quad \frac{\partial w^{out}}{\partial m_i} > 0,$$

so that both stronger outside opportunities (ψ_i) and greater mobility (m_i) raise counterfactual earnings. The proportional return to AI-sector entry relative to the outside option is then

$$R_i(A) = \frac{g(S_i(A), m_i)}{w^{out}(\psi_i, m_i)}.$$

Differentiating yields

$$\frac{\partial R_i}{\partial A} = \frac{g_S(S_i(A), m_i)\ f_A(\theta_i, A)}{w^{out}(\psi_i, m_i)},$$

and therefore

$$\frac{\partial^2 R_i}{\partial A \partial m_i} = \frac{f_A(\theta_i, A)}{\left[w^{out}(\psi_i, m_i)\right]^2} \left[g_{Sm}\left(S_i(A), m_i\right) \cdot w^{out}(\psi_i, m_i) - g_S\left(S_i(A), m_i\right) \cdot w_m^{out}(\psi_i, m_i)\right].$$

Because $f_A(\theta_i, A) > 0$ by assumption, the sign of $\partial^2 R_i / (\partial A \, \partial m_i)$ is governed by the bracketed term. In particular, it is positive if $g_{Sm}\left(S_i(A), m_i\right) \cdot w^{out}(\psi_i, m_i) > g_S\left(S_i(A), m_i\right) \cdot w_m^{out}(\psi_i, m_i)$, which can be written equivalently as a comparison of proportional sensitivities to market access:

$$\underbrace{\frac{\partial \ln g_S(S_i(A), m_i)}{\partial m_i}}_{\text{market-access amplification of the marginal AI payoff}} > \underbrace{\frac{\partial \ln w^{out}(\psi_i, m_i)}{\partial m_i}}_{\text{market-access effect on the outside option}}.$$

The left-hand side captures the proportional responsiveness of the marginal AI-sector return to skills to an increase in market access; the right-hand side captures the proportional responsiveness of the outside-option wage to m_i. This condition is economically plausible in our setting. The AI-sector payoff to skills is likely to be more market-access elastic because it is mediated by scale and matching in ways that are especially sensitive to connectivity. When m_i is high, graduates can access thicker AI labor markets, more specialized roles, and a wider set of firms that value AI-relevant skills; firms, in turn, can more easily recruit and deploy talent into productivity-enhancing tasks, steepening the skill premium in AI-intensive jobs. By contrast, the outside option is typically less skill-complementary with AI training and more closely tied to local or institution-specific demand, so improved access raises counterfactual wages by less. As a result, increases in m_i tend to amplify $g_S(S_i(A), m_i)$ more than they raise $w^{out}(\psi_i, m_i)$, making the inequality above a natural benchmark and preserving the implication that $\frac{\partial R_i}{\partial A}$ is increasing in market access even when w^{out} depends on m_i.