Paychecks Over Diplomas: How AI in Production Reduces Educational Investment

Santiago Caicedo* Gustavo de Souza†

March 26, 2026

Abstract

How does artificial intelligence (AI) affect educational investment and human capital accumulation? To answer this question, we use novel administrative software registry data from Brazil to measure regional AI exposure. Using as instrument shocks to AI development cost, we find that AI increases employment and wages, especially for young, less-educated workers in production occupations. These labor market opportunities pull high-school students from traditional schooling into work-compatible tracks: they move to night school, favor short vocational programs, or drop out of school entirely. These choices worsen educational outcomes, leading to higher failure rates, age-grade distortion, and lower test scores. We develop a task-based model showing that these patterns arise when AI complements low-skilled workers in production tasks, raising the opportunity cost of schooling and reducing overall human capital accumulation.

Key Words: Artificial Intelligence, Human Capital, Educational Choice, Automation
JEL: J24, O33, I25, J23, O15

*Caicedo: Northeastern University, sa.caicedo@northeastern.edu.
†de Souza: Federal Reserve Bank of Chicago, gustavo@microtomacro.net.

"If men learn this, it will implant forgetfulness in their souls; they will cease to exercise memory because they rely on that which is written."

—Socrates

1 Introduction

Technological progress changes the value of skills and, in doing so, the incentives to invest in human capital. Economists describe this relationship as a "race between education and technology," in which new technologies raise the demand for particular skills while schools expand their supply (Goldin and Katz 2008). During much of the twentieth century, technology ran ahead of education: successive waves of computerization and automation increased the demand for educated workers, increasing the skill-premium and the returns to schooling (Goldin and Katz 2008, Katz and Murphy 1992, Autor et al. 1998, Acemoglu 2002, Acemoglu and Restrepo 2020, de Souza and Li 2023).

Artificial intelligence has the potential to change the direction of this race (Autor 2024, Acemoglu and Restrepo 2020). Unlike earlier waves of automation that primarily displaced routine tasks (Autor et al. 2003, Autor and Dorn 2013), modern AI can also do or support non-routine cognitive work long done by educated, white-collar workers (Humlum and Vestergaard 2025, Eloundou et al. 2024, Gmyrek et al. 2023, Webb 2020, Felten et al. 2018). At the same time, AI can make expert knowledge cheaper and widely available, supporting younger and less experienced workers (de Souza 2025, Brynjolfsson et al. 2025, Noy and Zhang 2023, Dell'Acqua et al. 2023). If these changes raise demand for low-skilled work, they may decrease incentives to invest in education and reduce human capital accumulation.

Yet AI may have the opposite effect by automating less-skilled white-collar jobs. If AI replaces tasks done by junior workers, firms may hire fewer entry-level workers and increase the demand for experienced, high-skilled workers (Ide 2025, Hosseini Maasoum and Lichtinger 2025). Understanding which of these forces (complementarity or substitution) dominates is important because it determines how AI affects wage inequality and long-run growth through its impact on human capital accumulation.

We study these effects by developing a task-based model and testing its predictions using administrative data from Brazil. The model builds on the framework of Acemoglu and

Restrepo (2018, 2022), extended to include an educational choice margin. The key assumption is that AI complements low-skilled workers more than high-skilled workers in production tasks. AI software handles the cognitively demanding aspects of operating machinery, such as monitoring sensors, diagnosing problems, and adjusting settings. This enables less-educated workers to perform tasks that previously required expertise. The model yields three predictions: AI compresses wages by raising dropout wages relative to graduate wages, reallocates labor toward production activities, and reduces educational investment by raising the opportunity cost of schooling.

To test these predictions, we link administrative data on AI development to regional labor-market and education outcomes in Brazil. Our AI data come from the National Institute of Industrial Property (NIIP). In 1987, to support the emerging domestic IT sector, the Brazilian government created special copyright protection for source code registered with the NIIP. Registration later became standard industry practice: about 96% of software firms have at least one program registered, and industry experts stress the importance of this practice in any commercial software activity (de Souza 2025). We collect data on the universe of AI programs registered with the NIIP, including information such as the program's title, intended use, and technological features. We then link these records to administrative labor and education outcomes, letting us track how AI affects the labor market and schooling choices.

We build on de Souza (2025), who first assembled and validated the NIIP software registry data. That paper documents the institutional background of the registration system, establishes its coverage and representativeness, and uses it to study how AI affects occupational wages. Our analysis differs in two important respects. First, we study regional labor markets rather than occupations, allowing us to capture spillovers across occupations within local economies and to examine how AI reshapes the overall structure of employment in affected areas. Second, and more importantly, we study a new outcome that de Souza (2025) does not consider: educational investment. Our contribution is to show that AI's labor market effects have consequences for human capital accumulation, as students respond to changing labor market opportunities by adjusting their schooling decisions.

Our empirical strategy begins by building a measure of regional exposure to AI. For each

occupation, we use natural language processing to measure the similarity between official occupational task statements and the text descriptions of AI software in the NIIP registry. This strategy yields an occupation-level index of overlap between workers' tasks and AI capabilities, similar to other approaches in the literature (Webb 2020, de Souza and Li 2023, Kogan et al. 2023, Hampole et al. 2025, de Souza 2025). We then aggregate occupation exposure to the microregion level, weighting by each microregion's baseline occupational shares. Consequently, as the stock of AI software increases, microregions specialized in occupations with tasks closer to AI see larger increases in exposure than other regions.

To estimate the causal effect of AI, we use an instrumental-variable strategy that exploits exogenous variation in the cost of developing AI software. Our approach exploits the expansion of programming language ecosystems in the United States. As U.S. developers create more libraries, tutorials, and community support, they lower development costs globally by allowing others to use these resources (Delorey et al. 2007, Krein et al. 2010, Tambad et al. 2020, Vijayaraghavan et al. 2022). Because programming languages tend to specialize in particular application domains due to design and legacy (e.g., R for statistical work, COBOL for legacy business systems) (Nanz and Furia 2015, Rabah et al. 2010, Meyerovich and Rabkin 2013), these U.S.-driven expansions in resources lower the cost of producing some types of AI applications more than others. Consequently, U.S.-driven programming language expansions lower the cost of producing specific AI applications, disproportionately affecting Brazilian regions that specialize in those tasks. Our key identifying assumption is that shocks to U.S. programming language popularity are uncorrelated with local labor market dynamics in particular regions of Brazil.

We find that AI increases employment and wages, especially for young, less-educated workers in production occupations. A one standard deviation increase in AI exposure raises regional employment by 5% and hourly wages by 6%. The employment gain is small for workers over 30, no more than 2 percent, but much larger for workers aged 16 to 25, about 10 percent. These labor market opportunities pull high-school students from traditional schooling into work-compatible tracks: they move to night school, favor short vocational programs, or drop out of school entirely. These choices worsen educational outcomes, leading to higher failure rates, age-grade distortion, and lower test scores. Overall, these results suggest that

AI raises the opportunity cost of schooling by expanding job opportunities for low-skilled workers, encouraging early workforce entry and reducing human capital accumulation.

Literature. This paper contributes to the literature on technology and labor markets. Goldin and Katz (2008) characterize the twentieth century as a "race between education and technology," showing wage inequality since the 1980s reflects stagnating education rather than accelerating technology. Autor et al. (2003) demonstrate that computerization substitutes for routine tasks while complementing nonroutine work. Acemoglu and Restrepo (2018) formalize the "race between man and machine" where automation displaces labor while new tasks reinstate demand; Acemoglu and Restrepo (2022) show automation accounts for 50 to 70 percent of U.S. wage changes since 1980. Recent work finds that robots reduce employment and wages in exposed local labor markets (Acemoglu and Restrepo 2020), while AI affects hiring in establishments with exposed tasks (Acemoglu et al. 2022). Experimental evidence suggests AI compresses productivity differences across workers: Brynjolfsson et al. (2025) find larger gains for novice customer-support agents, while Noy and Zhang (2023) and Dell'Acqua et al. (2023) report similar patterns for writers and consultants.

Despite extensive research on technology's demand-side effects, far less is known about how human capital investment responds to AI-driven changes in labor market opportunities. Classic models predict that individuals invest in education when the returns exceed the costs (Becker 1964, Mincer 1974). Goldin and Katz (2008) document that educational attainment historically expanded to meet rising skill demand, but this supply response has slowed. The literature on school-to-work transitions emphasizes tradeoffs between vocational training, which facilitates immediate employment, and general education, which provides adaptability over the lifecycle (Hanushek et al. 2017). If AI increases demand for less-educated workers in production roles, students face altered incentives that may shift them toward earlier labor market entry at the expense of academic human capital. Yet no prior work documents how AI adoption causally affects educational investment decisions.

Our paper makes several contributions. First, we develop a task-based model showing that AI can complement low-skilled workers by making machinery easier to operate, raising the opportunity cost of schooling and reducing human capital accumulation. Second, we

use administrative software registry data from Brazil (de Souza 2025) that directly measures AI adoption. While de Souza (2025) uses these data to study occupational wage effects, we extend the analysis to regional labor markets and to educational outcomes. Third, we instrument AI adoption using exogenous changes in programming language ecosystems, addressing endogeneity concerns in technology-labor studies. Fourth, we provide causal evidence that AI increases demand for young, less-educated workers through capital complementarity, a channel distinct from the productivity compression documented in experimental work. Fifth, and most importantly, we document a supply-side response: students in AI-exposed labor markets drop out earlier or shift toward work-compatible educational tracks. These findings reveal that AI may reduce long-run human capital accumulation with lasting consequences for the "race between education and technology."

The remainder of the paper proceeds as follows. Section 2 describes the data, including the NIIP software registry and administrative labor market and education records. Section 3 develops the task-based model and derives testable predictions. Section 4 presents the empirical strategy for measuring regional AI exposure and our instrumental variable based on programming language ecosystems. Section 5 reports the main results on employment, wages, and educational choices. Section 6 concludes.

2 Data

We link four administrative datasets for Brazil: (i) the NIIP software registry, which provides detailed information on AI software developed in the country; (ii) RAIS, a matched employer-employee dataset covering the universe of formal workers; (iii) the School Census, which tracks enrollment across all schools; and (iv) occupational task descriptions from the Ministry of Labor. To construct our instrument, we use data on U.S. programming language popularity and Stack Overflow posts.

2.1 Administrative Software Registry Data

Since 1987, Brazil has granted copyright protection to computer programs registered with the National Institute of Industrial Property (NIIP). Registration provides fifty years of pro-

tection and serves as legal proof of ownership, which firms use to enforce intellectual property rights, access government procurement, obtain R&D credits, and signal to investors.[1] Because registration is inexpensive ($34 per filing) and widely required, it has become standard practice: de Souza (2025) reports that 96% of software development firms have at least one program registered.

Each registration contains the software title, application domains (intended uses), and technical classifications (underlying technologies). Figure 1 shows an example. We focus on software classified under artificial intelligence (technical class IA01), which includes applications ranging from predictive maintenance systems to automated quality control. Appendix A provides additional details on the classification system and examples of AI software in the registry.

The NIIP registry covers domestically developed software, which accounts for the vast majority of software used in Brazil. Imported computer services represented only 9% of domestic consumption in 2015, reflecting Brazil's large IT industry, regulatory barriers to imports, and high service trade costs (Arora and Gambardella 2005, Junqueira Botelho et al. 2005, Ezell et al. 2013, Benz and Jaax 2022).[2] We use data from de Souza et al. (2024), covering 37,677 software registrations from 1987 to 2022.

2.2 Labor Market and Education Data

RAIS. We measure labor market outcomes using the *Relação Anual de Informações Sociais* (RAIS), Brazil's administrative matched employer-employee dataset. RAIS provides annual data on every formal worker, including occupation, wages, education, and microregion. We aggregate these data to measure regional employment, average wages, and workforce composition.

School Census. We measure educational outcomes using the *Censo Escolar da Educação Básica*, the Ministry of Education's annual census of all schools in Brazil. The census provides enrollment-level data, which we use to measure enrollment in night classes, technical-

[1] See de Souza (2025) for a detailed discussion of the institutional background and registration incentives.
[2] Fewer than 0.02% of IT firms were foreign-owned in 2016.

Figure 1: **Example of a Software Registration**



FEDERATIVE REPUBLIC OF BRAZIL
MINISTRY OF THE ECONOMY
NATIONAL INSTITUTE OF INDUSTRIAL PROPERTY
DIRECTORATE OF PATENTS, COMPUTER PROGRAMS, AND TOPOGRAPHIES OF INTEGRATED CIRCUITS

Certificate of Software Registration

Case Nº: **BR512021001026-4**

The National Institute of Industrial Property issues this certificate of registration of a computer program, valid for 50 years starting from January 1st following the date of May 3, 2021, in accordance with §2, Article 2 of Law No. 9,609 of February 19, 1998.

Title: iMachine AI: Online Predictive Maintenance System for Electric Machines Using Machine Learning Techniques

Publication Date: 05/03/2021

Creation Date: 04/28/2021

Owner(s): TECHPLUS AUTOMACAO

Author(s): SAMARONE RUAS

Programming Language: PYTHON

Application Domain: AD-06; FQ-04; IF-10; IN-02

Technical Class: AP-01; AT-05; IA-01; IT-01; TC-03

Hash Algorithm: SHA-512

Digital Hash Summary:
cec9fe6467e2da2a9cc51fad00cfaedf52adc26dd5f352d4f4ed7c8600e399415383cc
2d542eab52cf2a49bf66955b20317e
bf21bafa7273d7edcdb534767588

Issued on: 05/25/2021

Approved by:
Carlos Alexandre
Fernandes Silva Head
of DIPTO

Notes: This figure shows a translated certificate of software registration. The certificate lists application domains (e.g., factory planning, metrology) and technical classifications.

vocational programs, and adult education (EJA).

Occupational Tasks. We construct AI exposure measures using task descriptions from the Ministry of Labor, which links each six-digit CBO-2002 occupation to detailed task statements. The dataset covers 2,641 occupations with an average of 64 tasks each.

2.3 Instrument Data

Programming Language Popularity. We track the global supply of programming resources using the Tiobe Index, which ranks languages by Google search results (TIOBE 2025). These counts proxy for the size of each language's ecosystem—tutorials, libraries, and community support—and thus the relative cost of developing software in that language.

Stack Overflow. We link programming languages to software applications using posts from Stack Overflow, a question-and-answer platform for programmers. Each post includes tags indicating the programming language and application area, allowing us to measure the comparative advantage of each language for specific technical classes.

3 A Model of AI and Educational Choice

This section develops a model that generates testable predictions about how AI affects labor markets and schooling decisions. The key insight is that AI makes machinery easier to operate, lowering the skill threshold for production work. This compresses wages across skill groups and raises the opportunity cost of staying in school. We present the main economic forces here; formal derivations appear in Appendix B.

3.1 Environment

We build on the task-based framework of Acemoglu and Restrepo (2018, 2022), extended to include an educational choice margin.

Tasks and activities. A final good is produced by combining a continuum of tasks $x \in \mathcal{T}$. Tasks divide into two types: *production tasks* $\mathcal{T}^p$ (manufacturing, agriculture, construction) and *office tasks* $\mathcal{T}^o$ (administration, clerical work). Output aggregates across tasks with elasticity of substitution $\lambda > 1$, so that tasks are substitutes but not perfect substitutes.

Workers. There are two types of workers: high school graduates (H) and dropouts (D). Workers choose their education before entering the labor market.

Technology. The key distinction is how tasks are performed:

- *Office tasks* require only labor. Graduates are more productive than dropouts: $\psi_H(x) > \psi_D(x)$ for most office tasks.
- *Production tasks* combine labor with capital. Critically, AI technology ψ_a affects how productive this combination is and affects workers differently depending on their education.

For production tasks, output takes the form:

$$y_p(x) = \psi_a(x)^{\eta_e} \ell_e(x)^{\alpha} k(x)^{1-\alpha} \tag{1}$$

where ℓ_e is labor of type $e \in \{D, H\}$, k is capital, and ψ_a represents AI technology embodied in software. In our empirical setting, ψ_a corresponds to the stock of AI software registered with the NIIP that targets production activities: manufacturing execution systems that coordinate workflows, predictive maintenance tools that monitor equipment, and quality control software that automates inspection. The parameter η_e governs how much this software improves productivity for each worker type.

Core assumption: AI complements low-skill workers in production. We assume:

$$\eta_H < \eta_D < 1 \tag{2}$$

This inequality is the engine of our results. It says that AI software raises the productivity of dropouts *more* than graduates in production tasks.

The economic intuition is that AI software handles the cognitively demanding aspects of operating machinery: monitoring sensors, diagnosing problems, adjusting settings, interpreting output quality. Consider the industrial copilots and predictive maintenance systems described in Section 2. These tools provide real-time guidance to machine operators, flagging anomalies and suggesting corrective actions. Such functions were previously performed by experienced, educated workers who had accumulated tacit knowledge over years on the job. As AI software takes over these functions, the machinery becomes "easier to use," and the

gap between educated and less-educated operators narrows. This pattern is consistent with experimental evidence showing that AI tools boost productivity more for less-experienced workers (Brynjolfsson et al. 2025, Noy and Zhang 2023, Peng et al. 2023).

3.2 How AI Affects the Labor Market

Consider an improvement in AI technology: ψ_a increases for a subset of production tasks. Four forces shape the labor market response.

Task reallocation. When AI is primitive, production tasks require expertise. Graduates, with their greater training, have a comparative advantage and perform most production work. As AI improves, it handles the complex aspects, and dropouts become cost-competitive. Formally, dropouts are cheaper when:

$$\frac{w_D^\alpha}{\psi_a^{\eta_D}} < \frac{w_H^\alpha}{\psi_a^{\eta_H}} \quad \Leftrightarrow \quad \psi_a > \left(\frac{w_D}{w_H}\right)^{\frac{\alpha}{\eta_D - \eta_H}} \equiv \hat{\psi}_a^D \tag{3}$$

Above this threshold, firms substitute toward dropout labor in production. At even higher levels of AI, full automation becomes cost-effective. Figure 2 illustrates these three regions.

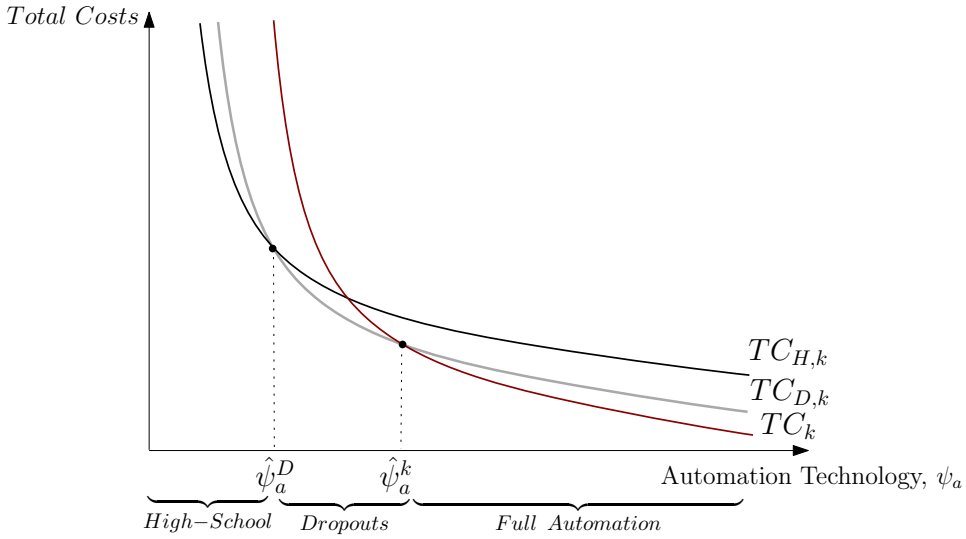
Capital complementarity. AI makes machinery more productive, raising the return to capital investment. Firms respond by deploying more equipment, which increases demand for workers who operate it. Since AI has made operation accessible to less-educated workers, this capital deepening disproportionately benefits dropouts.

Displacement. At sufficiently high AI levels, some tasks become fully automated, performed by capital alone with no labor input. This displaces workers from automated tasks.

Scale effects. Higher productivity expands output and raises demand for all inputs, including labor.

We show in Appendix B that, when graduates are more specialized in production tasks than dropouts and the AI shock is sufficiently large (Assumption 1), task reallocation and capital complementarity outweigh displacement for dropout workers:

Figure 2: Production Task Allocation and Automation Technology ψ_a



Notes: This figure illustrates how production tasks are allocated as AI technology ψ_a improves. The curves show the total cost of producing with graduates and capital ($TC_{H,k}$), dropouts and capital ($TC_{D,k}$), and full automation (TC_k). At low ψ_a, graduates operate machinery because $TC_{H,k}$ is lowest. At intermediate levels, dropouts become cost-competitive and take over. At high levels, full automation dominates. See Appendix B for derivations.

Proposition 1 (Wage Compression)**.** An AI shock that raises ψ_a for production tasks increases dropout wages relative to graduate wages: $d\ln w_D - d\ln w_H > 0$.

The intuition is that graduates lose on two margins: tasks reallocate away from them, and some of their tasks become automated. Dropouts, in contrast, gain from tasks reallocating toward them and from increased capital investment in production.

AI also affects the composition of employment:

Proposition 2 (Reallocation Toward Production)**.** An AI shock increases the share of dropout labor allocated to production relative to office tasks.

Two forces drive this prediction. First, AI raises productivity more in production (where it complements capital) than in office work. Second, the reduction in skill requirements specifically affects production tasks, drawing in less-educated workers who previously could not perform them.

3.3 How AI Affects Educational Investment

The labor market effects described above have direct implications for schooling decisions. Young individuals choose whether to complete high school by comparing the present value of wages against the costs of schooling. Let $V_e = w_e/\rho$ denote the present value of lifetime wages for a worker of type e, where ρ is the discount rate. The net benefit of completing high school is:

$$\Delta V = \frac{w_H - w_D}{\rho} - \tau \tag{4}$$

where τ captures schooling costs (time, effort, forgone earnings, direct expenses). Individuals differ in their idiosyncratic tastes or costs for education; those for whom ΔV exceeds their personal threshold complete high school.[3]

Proposition 1 established that AI compresses wages by raising w_D relative to w_H. This directly reduces ΔV:

Proposition 3 (Reduced Educational Investment)**.** If AI raises dropout wages relative to graduate wages, high school completion falls.

The intuition follows directly from the standard human capital model: when the wage premium from graduating shrinks, marginal students tip toward early labor market entry. The response is larger when the initial completion rate is near 50% (more students are marginal), when students are impatient (future wage gains matter less), and when preference heterogeneity is low (students respond similarly to incentives).

This prediction encompasses several observable behaviors: higher dropout rates, shifts toward work-compatible tracks (night school, vocational programs), and delayed completion as students interrupt their education to work and return later.

[3]There are many ways to model the educational choice margin, including dynamic models with sequential decisions, models with credit constraints, or models with uncertainty about future wages. Our qualitative results hold under any specification in which the supply of dropout labor is increasing in dropout wages relative to graduate wages.

3.4 Summary of Testable Predictions

Table 1 summarizes our model's predictions. The first two concern labor market outcomes; the third concerns educational responses.

Table 1: **Model Predictions**

	Prediction	Mechanism	Empirical Test
P1	Wage compression: AI raises dropout wages relative to graduate wages	Task reallocation and capital complementarity dominate displacement	Wage distribution by AI exposure
P2	Reallocation toward production, concentrated among less-educated workers	AI lowers skill threshold and raises productivity in production	Workforce composition by AI exposure
P3	Reduced educational attainment	Higher opportunity cost of schooling	Dropout rates, night school enrollment, vocational programs, test scores

Notes: This table summarizes the model's testable predictions, corresponding to Propositions 1–3.

Predictions P1 and P2 are consistent with evidence from de Souza (2025) at the occupation level. Our contribution is to test Prediction P3: the educational response to AI-driven changes in labor market opportunities.

4 Empirical Analysis

This section details our empirical strategy to test the model's predictions about AI's effects on labor markets and human capital. We start by building a measure of regional AI exposure from the text similarity between official occupational task descriptions and the universe of AI software registered in Brazil. We then estimate the effects of this exposure on labor and education outcomes in a panel model with fixed effects. To address endogeneity from local demand shocks, we use an instrumental-variables strategy based on supply-side shifts in AI development costs. The instrument uses the expansion of programming-language ecosystems in the United States and interacts this global resource growth with the technical requirements of specific AI applications to predict region-specific declines in software development costs.

4.1 AI Exposure

Text similarity between occupations and AI software. To measure an occupation's AI exposure, we link detailed task descriptions from the occupational classification system to the textual descriptions of AI software registered with the NIIP. This similarity provides a measure of the potential overlap between tasks workers perform and those AI could perform. We then exploit variation in the composition of workers across regions to measure how different regions are exposed to AI.

We construct the text-similarity measure in three steps. First, we collect the full set of task statements for every six-digit occupation from the *Classificação Brasileira de Ocupações* (CBO$\tilde{2}$002), provided by the Ministry of Labor. This set includes 49,660 different tasks, with each occupation performing 64 tasks on average. Second, for every AI software registered with the NIIP, we create a composite textual description by combining the software title and its registered application domains. This combination provides a list of tasks the software can perform.

Third, we calculate the text similarity between occupational tasks and AI software descriptions using natural language processing. We process the documents by removing stop words and lemmatizing terms, then use latent semantic analysis (LSA) to represent them in a high-dimensional semantic space. LSA maps synonyms and related terms (such as "inspection" and "quality control," or "prediction" and "forecasting") onto shared semantic dimensions, matching conceptually similar tasks even if their wording differs. Finally, for every occupation–software pair (o, s), we compute the cosine similarity between their descriptions, $\rho_{o,s}$. $\rho_{o,s}$ measures proximity between tasks performed by occupation o and those of the AI software s.

Regional Exposure to AI. Given that AI should disproportionately affect markets with a higher percentage of exposed occupations, we define the exposure of microregion m to AI

software in year t as[4]

$$AI\ Exposure_{m,t} = \sum_{o} \sum_{s} \lambda_{m,o,0} \times \rho_{o,s} \times \mathbb{I}_s \{t(s) \leq t\} \tag{5}$$

where $\lambda_{m,o,0}$ is the share of occupation o in region m's total employment in 2003, the first period of the sample. $\rho_{o,s}$ is the text-similarity between occupation o and software s. $\mathbb{I}_s \{t(s) \leq t\}$ is a dummy taking the value of one if software s was released by year t. According to the AI exposure in 5, a microregion m is more exposed to AI if it has a larger share of workers more exposed to AI.

4.2 Labor Market-level Specification

Main specification. To test the model's predictions about AI's labor market effects, we use the following specification:

$$y_{m,t} = \beta AI\ Exposure_{m,t} + \mu_m + \tilde{\mu} t + X' m, t\tilde{\theta} + \tilde{\epsilon}_{m,t} \tag{6}$$

where $y_{m,t}$ is an outcome of microregion m in year t, such as average wage, employment, or GDP. μ_m and $\tilde{\mu}_t$ are microregion and time fixed effects, respectively. As controls, $X_{m,t}$, we use the lag of the outcome variable, the regional exposure to non-AI software, and the regional exposure to AI patents in the U.S.

4.3 Instrument: AI Ease-of-Development

Endogeneity concerns: trends, directed technological change, and other regional shocks. An OLS estimate of β in the market-level regression (6) may be biased for three reasons. First, markets adopting AI earlier may differ systematically due to pre-existing economic or demographic trends. Second, AI firms may target regions or occupations where adoption is more profitable or technologically feasible, which could lead to a correlation between labor market trends and AI development. Third, local shocks, such as subsidies

[4]Each m represents a microregion, the Brazilian equivalent of a commuting zone in the United States. A microregion is defined by the Brazilian statistical agency as a set of economically connected municipalities. There are a total of 558 microregions.

or international trade exposure, may affect AI adoption and labor outcomes at the same time. To identify AI's causal effect, we instrument market-level AI exposure using exogenous variation in the cost of developing AI software, driven by changes in programming-language popularity outside of Brazil.

Exogenous variation: changes in the cost of creating AI software. The instrument builds on the idea that not all programming languages are equally suited for all applications, a point supported by a long literature in computer science (Delorey et al. 2007, Krein et al. 2010, Tambad et al. 2020, Vijayaraghavan et al. 2022). Each language holds a comparative advantage in specific tasks due to its design, history, and available libraries. For example, Python and R dominate data-intensive applications like statistics and machine learning, while C and C# are common in embedded systems and industrial automation.

Consequently, a rise in one language's popularity disproportionately lowers development costs for its best-suited applications. When such a language becomes more popular in the United States, which is the largest source of online programming resources, Brazilian developers in related domains benefit from lower costs when creating AI software in those same areas, leading to a larger supply of AI software. To illustrate how the instrument works, consider COBOL and R. COBOL (*Common Business-Oriented Language*), created in the 1950s, remains prevalent in legacy financial and administrative systems. Roughly 80% of in-person banking transactions and 95% of ATM operations worldwide use software built in COBOL (Hartman 2015). Despite this ubiquity, its global popularity has declined sharply, falling from the top 12 most-used languages in 2003 to 22nd in 2023, according to the Tiobe Index (TIOBE 2025). As the active COBOL programmer base shrinks, documentation and online resources become scarce, increasing the cost of maintaining or expanding AI applications for financial or administrative tasks. In contrast, R, a language for statistics and data analysis, has seen a surge in popularity and community support. Its extensive libraries and active user base make it highly efficient for developing analytical applications (Nanz and Furia 2015, Rabah et al. 2010).

These diverging trends mean, for instance, that it is much easier today to develop an AI tool for data analytics (using R) than one for banking automation (relying on COBOL). As

a consequence, regions whose occupational structures overlap more with data analysis experience larger increases in AI exposure, while those related to legacy administrative systems experience smaller ones. The identification strategy exploits this differential exposure: by comparing markets linked to programming languages that became cheaper to use with those tied to languages that became more costly, we isolate exogenous variation in AI diffusion unrelated to local labor-market trends.

Instrument definition. We first define the instrument at the occupation level and then aggregate it to the market level using employment shares. The instrument captures the ease of developing AI software targeting the tasks performed by each occupation. Formally, the ease of development for occupation o in year t is defined as:

$$\textit{AI Ease-of-Dev}_{o,t} = \sum_{l,p} \frac{\textit{Resources in US}_{l,t-2} \times \textit{Comparative Advantage}_{l,p} \times \overline{\textit{AI Exposure}}_{p,o,t-2}}{N_l N_p}, \tag{7}$$

where l indexes programming languages and p indexes software technical classes in the NIIP classification. The instrument aggregates three sources of information: (i) the time-varying availability of programming resources in the United States ($\textit{Resources in US}_{l,t-2}$), (ii) the comparative advantage of each programming language for specific technical classes ($\textit{Comparative Advantage}_{l,p}$), and (iii) the pre-existing relevance of those technical classes for each occupation ($\overline{\textit{AI Exposure}}_{p,o,t-2}$). All components are lagged by two years to ensure that they are predetermined relative to the observed labor-market outcomes. In the following paragraphs, we explain in detail each element of the instrument.

Programming-language resources in the United States. The first term measures the stock of publicly available programming resources for each language l in the United States, $\textit{Resources in US}_{l,t-2}$. We proxy this with the number of Google search hits for each language, as compiled by the Tiobe Index (TIOBE 2025). This variable captures the amount of online documentation, tutorials, libraries, and forums that developers can access. A large body of evidence shows that these resources substantially reduce development costs and increase

programmer productivity by facilitating code reuse and problem solving (Delorey et al. 2007, Krein et al. 2010, Tambad et al. 2020, Vijayaraghavan et al. 2022). Because Brazilian programmers frequently rely on global online communities, changes in U.S. programming-language ecosystems change the relative cost of developing AI software in Brazil, providing an exogenous source of variation. We lag this variable by two years to reflect the resources available at the time developers begin new projects.

Comparative advantage of programming languages. The second term captures the extent to which each programming language l is suited for applications in software technical class p: $\textit{Comparative Advantage}_{l,p}$. We estimate this measure using data from Stack Overflow, where each post is tagged with the programming language and application area discussed. The comparative advantage is defined as:

$$\textit{Comparative Advantage}_{l,p} = \frac{\textit{Stack Overflow posts tagging language } l \textit{ and technical class } p}{\textit{Stack Overflow posts tagging language } l} \tag{8}$$

This ratio reflects both intrinsic language characteristics and community specialization. For example, C++ and C# dominate game development (Wijdan 2020), while Python and Perl are widely used in bioinformatics (Fourment and Gillings 2008). The resulting $\textit{Comparative Advantage}_{l,p}$ variable links programming languages to the types of AI applications they are typically used for, allowing the instrument to capture how global programming trends shift the cost of developing software across particular technical domains.

Relevance of technical classes to occupations. The third term, $\overline{\textit{AI Exposure}}_{p,o,t-2}$, links technical classes to occupations. It measures the average textual similarity between AI software in technical class p and the tasks performed by occupation o, using only software registered before year $t-2$. This measure captures the baseline relevance of each technical class to each occupation. For instance, accountants are more exposed to AI software in resource accounting classes, while factory operators are more exposed to AI in automation systems. By using lagged exposure, we ensure that the weights reflect pre-existing technological relationships rather than contemporaneous adoption patterns, maintaining the

instrument's exogeneity.

Instrument aggregated at the regional level. As before, we aggregate the instrument at the region level and over time:

$$AI\ Ease\text{-}of\text{-}Dev_{m,t} = \sum_{t' \leq t} \sum_{o} \lambda_{m,o,0} \times AI\ Ease\text{-}of\text{-}Dev_{o,t'} \tag{9}$$

5 Empirical Results

We now test the model's predictions (summarized in Table 1) using our instrumental variables strategy. Recall that Proposition 1 predicts wage compression, Proposition 2 predicts reallocation toward production, and Proposition 3 predicts reduced educational attainment.

Strong first stage. The AI ease-of-development instrument strongly predicts regional AI exposure. As shown in Table 2, a one standard deviation increase in the AI ease-of-development instrument increases AI exposure by 1.5 standard deviations in the baseline specification. The corresponding F-statistics, all well above 30, confirm the instrument easily passes conventional weak-instrument thresholds and indicate a strong first stage.

Table 2: **AI Ease-of-Development Has a Strong Effect on Regional AI Exposure**

	(1)	(2)
	AI Exposure	*AI Exposure*
AI Ease-of-Dev	1.089***	1.548***
	(0.0181)	(0.0226)
Statistics:		
N	11,160	8,928
R^2	0.660	0.731
F	3614.2	1427.8

Notes: This table presents the estimates of the first stage. The instrument is given by equation 9 and the endogenous variable is the exposure to AI, defined in equation 5. All columns include microregion and year fixed effects. Column 2 adds controls for lagged outcomes, non-AI software exposure, and US AI patent exposure. Standard errors, clustered at the microregion level, are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

AI increases employment and wages. Table 3 reports the impact of AI on employment and wages. A one standard deviation increase in regional AI exposure raises local employment by roughly 5% and increases hourly wages by about 6%.

Table 3: **AI Increases Employment and Wages**

	(1) *log(Employment)*	(2) *log(Avg. Wage)*	(3) *log(Hourly Wage)*
AI Exposure	0.0554***	0.00940	0.0602***
	(0.0163)	(0.0117)	(0.0232)
N	8,928	8,928	8,928

Notes: This table reports estimates of the effect of AI on labor market outcomes, following the specification in Section 4. Column 1 reports the effect of AI on log employment; column 2 on the log of the average wage; and column 3 on the log of the hourly wage. AI exposure is defined in Equation 5 and is instrumented using the AI ease-of-development instrument defined in Equation 9. Controls include microregion fixed effects, year fixed effects, and the baseline controls described in Section 4. Standard errors, clustered at the microregion level, appear in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

As derived in Section 3, AI affects employment through three channels: a scale effect, capital complementarity, and displacement. The scale effect comes from AI increasing productivity, which expands output and raises demand for all inputs, including labor. Capital complementarity arises because AI makes machines more productive, increasing the demand for workers who operate them. Displacement works in the opposite direction, as AI fully automates tasks previously performed by labor. Taken together, the estimates in Table 3 indicate that the scale and capital complementarity effects dominate displacement, resulting in net employment growth.

The effect on average wages is also theoretically ambiguous. AI may raise wages by increasing labor demand, yet it could also lower them by changing workforce composition, for instance by drawing in more low-skilled workers. Table 3 shows that, at the regional level, the demand effect dominates, and AI exposure leads to higher average wages.

AI increases employment at production roles. Table 4 shows how AI exposure affects the composition of local employment across major occupational groups, providing evidence for Proposition 2. Column (1) shows that AI increases employment among production workers, a broad category that includes construction, manufacturing, mining, and agricultural

occupations. Columns (2) and (3) separate these effects and show that employment rises for both agricultural and manufacturing workers. Column (4) indicates that employment of managers also increases, likely reflecting the greater supervisory and coordination needs that accompany the expansion in production workers.

Table 4: **AI Increases Employment of Production Workers and Managers**

	(1)	(2)	(3)	(4)	(5)
	log(Production Workers)	*log(Agricultural Workers)*	*log(Manuf. Workers)*	*log(Managers)*	*log(Office Workers)*
AI Exposure	0.0927***	0.0950**	0.119***	0.151***	0.020
	(0.0245)	(0.0378)	(0.0246)	(0.0319)	(0.0253)
N	8,928	8,872	8,928	8,926	8,928

Notes: This table reports estimates of the effect of AI on employment by broad occupational group, following the specification in Section 4. Column 1 reports the effect of AI on log employment of production workers, which corresponds to CBO02 broad groups 3, 6, 7, 8 and 9; column 2 on log employment of agricultural workers, which corresponds to CBO02 broad group 6; column 3 on log employment of manufacturing workers, which corresponds to CBO02 broad groups 7, 8, and 9; column 4 on log employment of managers, which corresponds to CBO02 broad group 1; and column 5 on log employment of office and administrative workers, which corresponds to CBO02 broad group 4. AI exposure is defined in Equation 5 and is instrumented using the AI ease-of-development instrument defined in Equation 9. Controls include microregion fixed effects, year fixed effects, and the baseline controls described in Section 4. Standard errors, clustered at the microregion level, appear in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Column (5) shows the only group whose employment does not grow: office and administrative workers. This zero effect suggests the scale and substitution channels offset each other for these white-collar occupations. While AI automates some routine administrative tasks, it simultaneously increases demand for the non-automated tasks these workers still perform, leaving their net employment unchanged. This pattern is consistent with the model's prediction that AI primarily affects production tasks.

AI increases imports of machines. Table 5 shows that regions more exposed to AI significantly increase their imports of capital and high-technology machinery, consistent with the capital complementarity mechanism in the model. Column (1) documents that AI exposure raises total imports at the microregion level. Columns (2)–(4) reveal this increase comes from firms purchasing more foreign intermediate inputs and capital goods, not from households importing more final consumption goods.

Columns (5) and (6) decompose the growth in capital imports. Column (5) shows that AI exposure raises imports of machinery overall, while Column (6) finds an even stronger effect

on high-technology machines. These results align with equation (1), where AI technology ψ_a complements physical capital in production tasks.

Finally, Column (7) reports the effect of AI on the average unit value of imported goods, a proxy for input quality (Verhoogen 2008, Kugler and Verhoogen 2012). AI exposure increases the average import price, suggesting firms upgrade their input quality and adopt more advanced technologies as they incorporate AI into their production processes.

Table 5: **Effect of AI Exposure on Import Outcomes**

	(1) *log(Imports)*	(2) *log(Input Imports)*	(3) *log(Capital-Goods Imports)*	(4) *log(Final-Goods Imports)*	(5) *log(Imported Machinery)*	(6) *log(Imported High-Tech Machinery)*	(7) *log(Avg. Import Price)*
AI Exposure	1.340**	1.421**	1.179**	0.418	1.309**	1.204**	0.593**
	(0.584)	(0.569)	(0.597)	(0.528)	(0.616)	(0.580)	(0.279)
N	8,928	8,928	8,928	8,928	8,928	8,928	5,415

Notes: This table reports estimates of the effect of AI exposure on import outcomes, following the specification in Section 4. Column 1 reports the effect of AI on log total imports; column 2 on log intermediate-input imports; column 3 on log capital-goods imports; column 4 on log final-goods imports; column 5 on log imported machinery; column 6 on log imported high-tech machinery; and column 7 on log average import unit values. The classification of goods comes from SECEX, and machinery is defined as HS codes 8401–9209. The definition of high-tech machinery comes from the U.S. Census advanced-technology-products list. Because these import measures can be zero, we add one before taking logs. Standard errors, clustered at the microregion level, appear in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

AI increases employment of young workers who have not finished high school. The impact of AI on local employment is concentrated among younger and less educated workers, consistent with Proposition 1's prediction of wage compression. Table 6 reports how AI changes the composition of the workforce. Column (1) shows that AI reduces the average years of education among employed workers, while Column (2) shows a decline in average experience. Column (3) shows that AI increases the share of high-school dropouts in the workforce. Taken together, these results indicate that regions more exposed to AI hire disproportionately more young workers who have not completed high school. This pattern is consistent with the model's key assumption (inequality 2): AI makes machinery easier to operate, raises the productivity of less-experienced workers, and expands the set of tasks they can perform, leading firms to increase hiring among these groups.

Figure 3 shows how the employment effects of AI vary according to the age of the worker. The impact is largest for workers of high-school age: a one standard deviation increase in AI

exposure raises employment among 16- to 18-year-olds by nearly 20%. The effects gradually decline with age, remaining positive for workers up to their late twenties and approaching zero for older workers. This pattern is consistent with AI increasing the demand for younger, less-experienced workers who can more easily complement new technologies in production (de Souza 2025, Brynjolfsson et al. 2025, Gruber et al. 2020, Choi et al. 2023, Dell'Acqua et al. 2023, Noy and Zhang 2023, Peng et al. 2023).

Table 6: **Effect of AI Exposure on Worker Characteristics**

	(1) *log(Average Years of Education)*	(2) *log(Average Experience)*	(3) *Share High-School Dropouts*
AI Exposure	-0.0160^{***}	-0.0263^{***}	0.0302^{***}
	(0.00279)	(0.00994)	(0.00468)
N	8,928	8,928	8,928

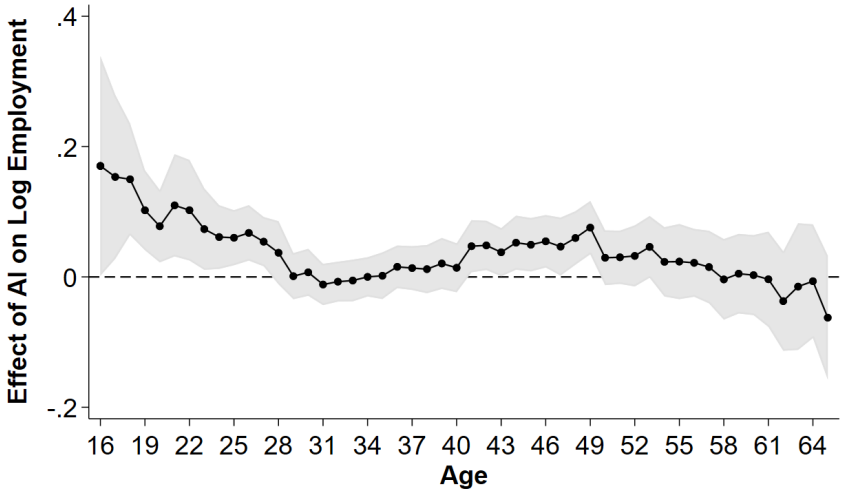
Notes: This table reports estimates of the effect of AI on worker characteristics, following the specification in Section 4. Column 1 reports the effect of AI on log average years of education; column 2 on log average experience; and column 3 on the share of high-school dropouts. AI exposure is defined in Equation 5 and is instrumented using the AI ease-of-development instrument defined in Equation 9. Controls include microregion fixed effects, year fixed effects, and the baseline controls described in Section 4. Standard errors, clustered at the microregion level, appear in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

These young workers who enter the labor market in response to the AI shock are also more likely to forego completing high school. Figure 4 plots the effect of AI exposure on the share of workers who have not finished high school by age group. AI sharply increases the share of young workers without a high-school degree. For workers under age 18, part of this increase is mechanical, as many have not yet had the opportunity to graduate. But the fact that the effect persists into later ages indicates a behavioral response: students are permanently dropping out of high school or delaying their graduation in order to join the labor market when AI-driven employment opportunities expand.

Students choose educational paths to conciliate work and school. Table 7 provides direct evidence for Proposition 3, showing how AI exposure affects students' educational choices. Column (1) indicates that regions more exposed to AI experience an increase in the number of night-school students. In Brazil, night shifts in basic education are typically chosen by students who work during the day and attend school in the evening. This pattern

reinforces the earlier evidence that AI draws young individuals into the labor market and shows how they adapt their schooling decisions: rather than leaving education entirely, many adjust by shifting to night school so they can balance work and study.

Figure 3: **AI Increases Employment of High-School-Age Workers**



Notes: This figure plots the estimated effect of AI exposure on log employment by worker age group, following the specification in Section 4. Each point corresponds to a separate 2SLS estimate for a given age, where AI exposure is defined in Equation 5 and instrumented using the AI ease-of-development instrument defined in Equation 9. All regressions include microregion fixed effects, year fixed effects, and the baseline controls described in Section 4. Error bars report 95% confidence intervals based on standard errors clustered at the microregion level.

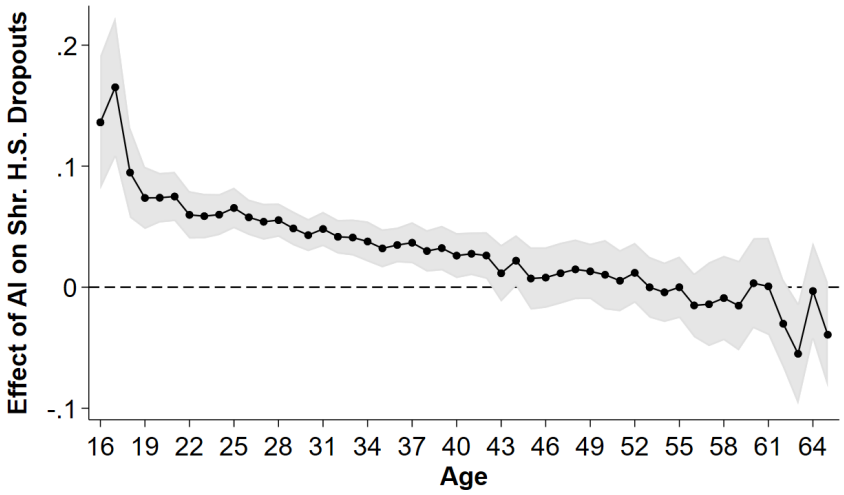
Column 2 shows that AI exposure also increases the number of students enrolled in technical–professional high-school programs. In Brazil, these programs combine regular high-school coursework with vocational training and typically require students to complete an apprenticeship in a firm, preparing them for technical occupations immediately after graduation. Because these courses are closely tied to shop-floor and production roles, the increase in enrollment suggests that AI leads students to select shorter, more applied educational paths that facilitate a faster transition into production-related jobs.

Column 3 of Table 7 shows that AI raises high-school grade repetition. As AI draws more students into work, they must juggle jobs and school; it is reasonable that this leads some to fall behind in their schoolwork. As a result, the share of students who fail a grade increases.

Consistent with the rise in grade repetition, column 4 of Table 7 shows that AI raises enrollment in adult-education programs. In Brazil, these programs serve working adults over

18 who lack a high-school diploma. An increase in enrollment in these programs can come from more students repeating grades or from students leaving high school and completing it later as adults.

Figure 4: **AI Increases Employment of H.S. Dropouts**



Notes: This figure plots the estimated effect of AI exposure on the share of high-school dropouts by worker age group, following the specification in Section 4. Each point corresponds to a separate 2SLS estimate for a given age, where AI exposure is defined in Equation 5 and instrumented using the AI ease-of-development instrument defined in Equation 9. All regressions include microregion fixed effects, year fixed effects, and the baseline controls described in Section 4. Error bars report 95% confidence intervals based on standard errors clustered at the microregion level.

Table 7: **Effect of AI Exposure on School Enrollment**

	(1) *log(Night-School Students)*	(2) *log(Students in Technical–Vocational H.S.)*	(3) *H.S. Failure Rate*	(4) *log(Adult-Education Students)*
AI Exposure	0.0674***	0.136**	0.0119**	0.0915***
	(0.0191)	(0.0552)	(0.00486)	(0.0319)
N	10,039	8,722	8,927	10,025

Notes: This table reports estimates of the effect of AI exposure on education outcomes, following the specification in Section 4. Column 1 reports the effect of AI on the log number of night-school students; column 2 on the log number of students enrolled in technical-vocational high-school programs; column 3 on the share of high-school students who failed a grade; and column 4 on the log number of students enrolled in adult-education (EJA) programs. AI exposure is defined in Equation 5 and is instrumented using the AI ease-of-development instrument defined in Equation 9. Controls include microregion fixed effects, year fixed effects, and the baseline controls described in Section 4. Standard errors, clustered at the microregion level, appear in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Work-compatible schooling reduces learning. As students choose tracks that accommodate a work schedule, their academic learning falls, as reflected in ENEM scores, Brazil's

standardized national college entrance exam. Table 8 reports the effect of AI exposure on average exam performance. AI lowers the overall score and the score in most areas. A one standard deviation increase in AI exposure reduces the overall score by about 0.05 standard deviations, with larger declines in languages, human sciences, and math.

There are two mechanisms that could explain the decline in the entrance exam scores, both consistent with the model. First, as AI-driven labor demand opens more production-job opportunities, some students may see less value in university and put in less effort to prepare for the exam. Second, because many of these students are balancing work with school and are shifting into more flexible or shorter educational tracks, they may have learned less during high school, resulting in weaker performance on the exam. Regardless of the precise mechanism, Table 8 shows that even students who finish high school learn less in regions where AI raises demand for young workers.

Table 8: **Effect of AI on College Entrance Exam Scores**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	College Entrance Exam Score (ENEM)						
	Overall	*Multiple-Choice*	*Writing*	*Natural Sciences*	*Human Sciences*	*Languages*	*Math*
AI Exposure	-0.0514^{*}	-0.0261^{**}	-0.0464	0.0124	-0.0974^{***}	-0.162^{***}	-0.0853^{**}
	(0.0275)	(0.0112)	(0.0455)	(0.0292)	(0.0270)	(0.0369)	(0.0365)
N	5,580	5,580	5,580	5,580	5,580	5,580	5,580

Notes: This table reports estimates of the effect of AI exposure on college-entrance exam performance (ENEM), following the specification in Section 4. The ENEM exam has two broad components: the multiple choice part and the writing part. The multiple choice tests on natural sciences, human sciences, languages, and mathematics. Column 1 reports the effect of AI on the overall ENEM score; column 2 on the multiple-choice score; column 3 on the writing score; column 4 on the natural-sciences score; column 5 on the human-sciences score; column 6 on the languages score; and column 7 on the math score. AI exposure is defined in Equation 5 and is instrumented using the AI ease-of-development instrument defined in Equation 9. Controls include microregion fixed effects, year fixed effects, and the baseline controls described in Section 4. Standard errors, clustered at the microregion level, appear in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

AI reduces human capital by pulling high-school students into work. Our results confirm the model's three main predictions. First, consistent with Proposition 1, AI compresses wages by raising demand for young, low-skilled workers. Second, consistent with Proposition 2, AI reallocates labor toward production activities. Third, consistent with Proposition 3, AI reduces educational investment: more students leave high school, and those who stay learn less. The mechanism operates through the opportunity cost of school-

ing: AI makes low-skill work more attractive than staying in school and pulls young workers into the labor force earlier. In the race between education and technology, AI accelerates some workers' entry while slowing their skill development.

6 Conclusion

This paper studies how artificial intelligence affects human capital accumulation. Using administrative software registry data from Brazil and an instrumental variable strategy based on programming language ecosystems, we find that AI increases employment and wages, particularly for young, less-educated workers in production occupations. Students respond to these labor market opportunities by dropping out of high school or shifting toward work-compatible educational tracks. These choices reduce educational attainment and academic performance, as measured by grade repetition rates and college entrance exam scores. We develop a task-based model showing that AI complements low-skilled workers by making machinery easier to operate, raising the opportunity cost of schooling and reducing aggregate human capital accumulation.

These findings suggest a reversal in the "race between education and technology" that characterized much of the twentieth century. While earlier waves of automation increased the returns to education and expanded schooling, AI may move in the opposite direction by creating immediate employment opportunities that pull young workers from the classroom to the factory floor. Like Socrates' concern that writing would cause men to "cease to exercise memory," AI may reduce investment in foundational human capital by making certain skills appear less necessary.

Several questions remain for future research. First, what are the long-run consequences of reduced educational investment for productivity growth and wage inequality? Our estimates capture short-run responses, but the effects may compound as cohorts with less education move through the labor market. Second, how do these patterns differ across developing and advanced economies? Brazil's large informal sector and relatively low educational attainment may amplify the pull of AI-driven employment opportunities in ways that would not occur in countries with higher baseline education levels. Third, which types of education matter

most in an AI economy? If AI reduces the returns to routine cognitive skills while raising the returns to creativity, judgment, and interpersonal skills, educational systems may need fundamental restructuring rather than simple expansion. Finally, can policy interventions preserve educational investment while allowing workers to benefit from AI-driven opportunities? Understanding these tradeoffs is essential for ensuring that AI contributes to shared prosperity rather than entrenching inequality through reduced human capital accumulation.

References

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

Acemoglu, D., D. Autor, J. Hazell, and P. Restrepo (2022): "Artificial Intelligence and Jobs: Evidence from Online Vacancies," *Journal of Labor Economics*, 40, S293–S340.

Acemoglu, D. and P. Restrepo (2018): "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment," *American Economic Review*, 108, 1488–1542.

——— (2020): "Robots and Jobs: Evidence from US Labor Markets," *Journal of Political Economy*, 128, 2188–2244.

——— (2022): "Tasks, Automation, and the Rise in U.S. Wage Inequality," *Econometrica*, 90, 1973–2016.

Arora, A. and A. Gambardella (2005): "The Globalization of the Software Industry: Perspectives and Opportunities for Developed and Developing Countries," in *Innovation Policy and the Economy, Volume 5*, National Bureau of Economic Research, Inc, NBER Chapters, 1–32.

Associação Brasileira das Empresas de Software (ABES) (2024): "Por que registrar meu software?" Accessed: 2024-11-14.

Autor, D. (2024): "Applying AI to Rebuild Middle Class Jobs," NBER Working Paper No. 32140.

Autor, D. H. and D. Dorn (2013): "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103, 1553–1597.

Autor, D. H., L. F. Katz, and A. B. Krueger (1998): "Computing Inequality: Have Computers Changed the Labor Market?" *The Quarterly Journal of Economics*, 113, 1169–1213.

Autor, D. H., F. Levy, and R. J. Murnane (2003): "The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, 118, 1279–1333.

Becker, G. S. (1964): *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, Chicago, IL: University of Chicago Press.

Benz, S. and A. Jaax (2022): "The costs of regulatory barriers to trade in services: New estimates of ad valorem tariff equivalents," *Economics Letters*, 212, 110057.

Brynjolfsson, E., D. Li, and L. Raymond (2025): "Generative AI at Work," *Quarterly Journal of Economics*, 140, 889–942.

Choi, J. H., D. Schwarcz, and K. E. Yeh (2023): "AI Assistance in Legal Analysis: An Empirical Study," Legal Studies Research Paper 23-22, University of Minnesota Law School.

de Souza, G. (2025): "Artificial Intelligence in the Office and the Factory: Evidence from Administrative Software Registry Data," Chicago Fed Working Paper 2025-11, Federal Reserve Bank of Chicago.

de Souza, G., J. S. Herbstman, and J. Mannion (2024): "How Demand for New Skills Affects Wage Inequality: The Case of Software Programmers," Working Paper Series WP 2024-19, Federal Reserve Bank of Chicago.

de Souza, G. and H. Li (2023): "Robots, Tools, and Jobs: Evidence from Brazilian Labor Markets," Working Paper Series WP 2023-42, Federal Reserve Bank of Chicago.

Dell'Acqua, F., E. McFowland III, E. Mollick, H. Lifshitz-Assaf, K. Kellogg, S. Rajendran, L. Krayer, F. Candelon, and K. R. Lakhani (2023): "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality," Harvard Business School Working Paper No. 24-013.

Delorey, D., C. Knutson, and S. Chun (2007): "Do Programming Languages Affect Productivity? A Case Study Using Data from Open Source Projects," 8–8.

Eloundou, T., S. Manning, P. Mishkin, and D. Rock (2024): "GPTs are GPTs: Labor Market Impact Potential of Large Language Models," *Science*, 384, 1306–1308.

Ezell, S. J., R. D. Atkinson, and M. A. Wein (2013): "Localization Barriers to Trade: Threat to the Global Innovation Economy," Technical Report –, Information Technology & Innovation Foundation, Washington, DC, available at `https://www2.itif.org/2013-localization-barriers-to-trade.pdf`.

Felten, E., M. Raj, and R. Seamans (2018): "A Method to Link Advances in Artificial Intelligence to Occupational Abilities," in *AEA Papers and Proceedings*, vol. 108, 54–57.

Fourment, M. and M. R. Gillings (2008): "A comparison of common programming languages used in bioinformatics," *BMC Bioinformatics*, 9, 82.

Gmyrek, P., J. Berg, and D. Bescond (2023): "Generative AI and jobs a global analysis of potential effects on job quantity and quality," Tech. rep.

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

Gruber, J., B. R. Handel, S. H. Kina, and J. T. Kolstad (2020): "Managing Intelligence: Skilled Experts and AI in Markets for Complex Products," NBER Working Papers 27038, National Bureau of Economic Research, Inc.

Hampole, M., D. Papanikolaou, L. D. Schmidt, and B. Seegmiller (2025): "Artificial Intelligence and the Labor Market," NBER Working Papers 33509, National Bureau of Economic Research, Inc.

Hanushek, E. A., G. Schwerdt, L. Woessmann, and L. Zhang (2017): "General Education, Vocational Education, and Labor-Market Outcomes over the Lifecycle," *Journal of Human Resources*, 52, 48–87.

Hartman, T. (2015): "COBOL Blues," `https://www.reuters.com/graphics/USA-BANKS-COBOL/010040KH18J/`, accessed: 16 Feb 2025.

Hosseini Maasoum, S. M. and G. Lichtinger (2025): "Generative AI as Seniority-Biased Technological Change: Evidence from U.S. Résumé and Job Posting Data," Ssrn working paper, SSRN, posted 2025-09-08; last revised 2025-11-11.

Humlum, A. and E. Vestergaard (2025): "The Unequal Adoption of ChatGPT Exacerbates Existing Inequalities Among Workers," *Proceedings of the National Academy of Sciences*, 122.

Ide, E. (2025): "Automation, AI, and the Intergenerational Transmission of Knowledge," Papers 2507.16078, arXiv.org.

Junqueira Botelho, A. J., G. Stefanuto, and F. Veloso (2005): "The Brazilian Software Industry," in *From Underdogs to Tigers: The Rise and Growth of the Software Industry in Brazil, China, India, Ireland, and Israel*, Oxford University Press.

Katz, L. F. and K. M. Murphy (1992): "Changes in Relative Wages, 1963–1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 107, 35–78.

Kogan, L., D. Papanikolaou, L. D. Schmidt, and B. Seegmiller (2023): "Technology and Labor Displacement: Evidence from Linking Patents with Worker-Level Data," Working Paper 31846, National Bureau of Economic Research.

Krein, J., A. Maclean, C. Knutson, D. Delorey, and D. Eggett (2010): "Impact of Programming Language Fragmentation on Developer Productivity: A Sourceforge Empirical Study," *IJOSSP*, 2, 41–61.

Kugler, M. and E. Verhoogen (2012): "Prices, Plant Size, and Product Quality," *The Review of Economic Studies*, 79, 307–339.

Meyerovich, L. A. and A. S. Rabkin (2013): "Empirical analysis of programming language adoption," in *Proceedings of the 2013 ACM SIGPLAN International Conference on Object Oriented Programming Systems Languages & Applications*, New York, NY, USA: Association for Computing Machinery, OOPSLA '13, 1–18.

Mincer, J. (1974): *Schooling, Experience, and Earnings*, New York, NY: National Bureau of Economic Research.

Nanz, S. and C. A. Furia (2015): "A Comparative Study of Programming Languages in Rosetta Code," in *2015 IEEE/ACM 37th IEEE International Conference on Software Engineering*, IEEE, 778–788.

Noy, S. and W. Zhang (2023): "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence," *Science*, 381, 187–192.

Peng, S., E. Kalliamvakou, P. Cihon, and M. Demirer (2023): "The Impact of AI on Developer Productivity: Evidence from GitHub Copilot," .

Rabah, S., J. Li, M. Liu, and Y. Lai (2010): "Comparative Studies of 10 Programming Languages within 10 Diverse Criteria – a Team 7 COMP6411-S10 Term Report," .

Tambad, S., R. Nandwani, and S. K. McIntosh (2020): "Analyzing programming languages by community characteristics on Github and StackOverflow," .

TIOBE (2025): "TIOBE Programming Community Index," `https://www.tiobe.com/tiobe-index/`, accessed: 2025-07-18.

Verhoogen, E. A. (2008): "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector," *The Quarterly Journal of Economics*, 123, 489–530.

Vijayaraghavan, S., J. Song, T. Guan, S. Choi, and S. Kulkarni (2022): "Influence of Communication Among Shared Developers on the Productivity of Open Source Software Projects," .

Webb, M. (2020): "The Impact of Artificial Intelligence on the Labor Market," Stanford University Working Paper.

Wijdan, A. (2020): "Comparison of Programming Languages in Game Development," .

A Data Appendix

This appendix provides additional details on the software registry data and AI applications in Brazil.

A.1 Software Registration System

Brazilian software law (Law 7,646/1987, updated by Law 9,609/1998) defines software as a "set of instructions or statements, written in proper language, to be used directly or indirectly by a computer to obtain a certain result." Unlike patents, software copyright protects the source code itself rather than the underlying idea, and registration does not require novelty or an inventive step.

Firms register software for six main reasons:

1. *Intellectual protection*: Registration provides legal proof of ownership for enforcing copyright.
2. *Government procurement*: Certificates are required to bid on public IT contracts.
3. *Access to credit*: Registration serves as evidence of innovation for R&D loans.
4. *Investor signaling*: Most venture capital investors require registration of proprietary software (Associação Brasileira das Empresas de Software (ABES) 2024).
5. *Low cost*: Registration costs only $34 per filing.
6. *Developer credentials*: Programmers listed on registrations receive a permanent wage premium of approximately 5% (de Souza et al. 2024).

A.2 Classification System

Each software registration includes two types of classifications:

Application domains indicate the intended use of the software. There are 225 domains grouped into 35 broad categories, including administration, agriculture, manufacturing,

health, and education. For example, the software in Figure 1 is classified under factory planning (AD-06), metrology (FQ-04), information management (IF-10), and manufacturing technology (IN-02).

Technical classes describe the underlying technology. There are 97 classes grouped into 18 broad categories, including artificial intelligence (IA), automation systems (AT), and data management (GI). We identify AI software as those with technical class IA01.

A.3 Examples of AI Software

The registry contains diverse AI applications. Two examples illustrate the range:

Predictive maintenance (Imachine AI). Developed by Techplus and shown in Figure 1, this software connects to factory sensors to monitor equipment performance. Using machine learning, it predicts failures and schedules preventive maintenance, reducing unplanned downtime.

Quality control systems. Several firms have registered AI software for automated visual inspection. For example, Tbit's image-recognition system identifies impurities in products such as seeds and animal feed, flagging defects in real time and relaying quality metrics to manufacturing execution systems.

These examples reflect a broader pattern: approximately 44% of AI software in the registry targets production activities rather than administrative functions (de Souza 2025).

A.4 Limitations

The NIIP database has four limitations:

1. *Domestic coverage only*: The registry excludes foreign-developed software. However, imported software accounts for only 9% of domestic consumption.

2. *Legal rather than technical units*: A single registration may bundle multiple programs. We address this by analyzing textual descriptions rather than registration counts.

3. *Underrepresentation of mobile apps*: Developers on platforms with built-in IP protection (e.g., app stores) have weaker incentives to register. This is unlikely to affect our estimates, which focus on firm-facing software.

4. *Version updates*: Firms often register only initial versions, so late additions of AI features may go unrecorded. This would attenuate our estimates, making them conservative.

B Theoretical Appendix

This appendix provides the full model setup and proofs of the propositions in Section 3.

B.1 Full Model Setup

Aggregate production. Output combines production and office activities:

$$Y = F(y_p, y_o) = \left(y_p^{\frac{\varsigma-1}{\varsigma}} + y_o^{\frac{\varsigma-1}{\varsigma}} \right)^{\frac{\varsigma}{\varsigma-1}}$$

with elasticity $\varsigma > 1$. Within each activity, output aggregates over tasks:

$$y_s = \left(\int_{x \in \mathcal{T}^s} y_s(x)^{\frac{\lambda-1}{\lambda}} dx \right)^{\frac{\lambda}{\lambda-1}}, \quad s \in \{p, o\},$$

with $\lambda > 1$ so tasks are imperfect substitutes.

Office task production. Each office task is produced by one worker type:

$$y_o(x) = \psi_D(x)\ell_D(x) \quad \text{or} \quad y_o(x) = \psi_H(x)\ell_H(x)$$

Production task technology. Each production task can be produced through three methods:

$$y_p(x) = \begin{cases} \psi_a(x)^{\eta_D} \ell_D(x)^{\alpha} k(x)^{1-\alpha} & \text{(dropouts + capital)} \\ \psi_a(x)^{\eta_H} \ell_H(x)^{\alpha} k(x)^{1-\alpha} & \text{(graduates + capital)} \\ \psi_a(x) k(x) & \text{(full automation)} \end{cases}$$

where $\eta_H < \eta_D < 1$. Capital is produced from the final good at marginal cost $1/q(x)$.

Educational choice. Individual i receives utility:

$$U_D(i) = V_D + \varepsilon_{D,i}, \quad U_H(i) = V_H - \tau + \varepsilon_{H,i}$$

where $V_e = w_e/\rho$ represents the present value of lifetime wages, $\rho > 0$ is the discount rate, $\tau > 0$ captures the costs of high school completion, and $\varepsilon_{e,i} \sim$ Type I Extreme Value with scale parameter σ. The share completing high school is:

$$\pi_H = \frac{1}{1 + \exp(-\Delta V/\sigma)}, \quad \text{where} \quad \Delta V = \frac{w_H - w_D}{\rho} - \tau$$

Competitive equilibrium. A competitive equilibrium consists of wages $\{w_D, w_H\}$, task allocations $\{\mathcal{T}_D, \mathcal{T}_H, \mathcal{T}_{D,k}, \mathcal{T}_{H,k}, \mathcal{T}_k\}$, and factor allocations $\{\ell_D(x), \ell_H(x), k(x)\}$ such that: (i) firms allocate tasks to minimize costs, (ii) individuals choose education to maximize utility, and (iii) markets clear.[5]

B.2 Task Allocation

We first establish how tasks are allocated across worker types and automation as a function of the AI technology level.

Lemma B.1 (Task Allocation by Automation Level)**.** *Suppose* $\eta_H < \eta_D < 1$ *and* $\frac{w_D}{1/q} < \left(\frac{1-\alpha}{\alpha}\right)^{\frac{1-\alpha}{\alpha}}$. *Then there exist thresholds* $0 < \hat{\psi}_a^D < \hat{\psi}_a^k < \infty$ *such that production tasks are allocated as follows:*

[5]The existence of equilibrium in this task-based framework follows Acemoglu and Restrepo (2022).

- *Low automation ($\psi_a < \hat{\psi}_a^D$): performed by graduates with capital*
- *Medium automation ($\hat{\psi}_a^D \leq \psi_a < \hat{\psi}_a^k$): performed by dropouts with capital*
- *High automation ($\psi_a \geq \hat{\psi}_a^k$): fully automated*

where $\hat{\psi}_a^D = \left(\frac{w_D}{w_H}\right)^{\frac{\alpha}{\eta_D - \eta_H}}$ *and* $\hat{\psi}_a^k = \left(\frac{\alpha}{1-\alpha}\right)^{\frac{1-\alpha}{1-\eta_D}} (w_D q)^{\frac{\alpha}{1-\eta_D}}$.

Proof. We first establish the threshold $\hat{\psi}_a^D$. Production with dropouts is cheaper than with graduates when:

$$\frac{w_D^\alpha}{\psi_a^{\eta_D}} < \frac{w_H^\alpha}{\psi_a^{\eta_H}} \iff \left(\frac{w_D}{w_H}\right)^\alpha < \psi_a^{\eta_D - \eta_H} \iff \psi_a > \left(\frac{w_D}{w_H}\right)^{\frac{\alpha}{\eta_D - \eta_H}} = \hat{\psi}_a^D$$

Next, we establish the full automation threshold $\hat{\psi}_a^k$. Since dropouts are the lower-cost worker type for $\psi_a \geq \hat{\psi}_a^D$, full automation is cheaper than partial automation when:

$$\frac{1}{q\psi_a} < \frac{w_D^\alpha}{\psi_a^{\eta_D}} \left(\frac{\alpha}{1-\alpha}\right)^{1-\alpha} q^{-(1-\alpha)} \iff \psi_a > \left(\frac{\alpha}{1-\alpha}\right)^{\frac{1-\alpha}{1-\eta_D}} (w_D q)^{\frac{\alpha}{1-\eta_D}} = \hat{\psi}_a^k$$

Finally, we verify $\hat{\psi}_a^D < \hat{\psi}_a^k$. Since $w_D < w_H$, we have $\hat{\psi}_a^D < 1$. The inequality $\hat{\psi}_a^D < \hat{\psi}_a^k$ is equivalent to:

$$1 < \left(\frac{\alpha}{1-\alpha}\right)^{\frac{1-\alpha}{1-\eta_D}} (w_D q)^{\frac{\alpha}{1-\eta_D}} \iff \frac{w_D}{1/q} < \left(\frac{1-\alpha}{\alpha}\right)^{\frac{1-\alpha}{\alpha}}$$

which holds by assumption. Note that $\left(\frac{1-\alpha}{\alpha}\right)^{\frac{1-\alpha}{\alpha}} \geq e^{-1/e} \approx 0.692$ for all $\alpha \in (0,1)$. □

B.3 Wage Effects

Market clearing requires:

$$\Gamma_e w_e^{-\lambda} Y^{1/\varsigma} y_o^{(\varsigma-1)/\varsigma} + \Gamma_{a,e} w_e^{-\alpha\lambda-(1-\alpha)} Y^{1/\varsigma} y_p^{(\varsigma-1)/\varsigma} = L_e$$

where the productivity aggregates are defined as:

$$\Gamma_e = \int_{\mathcal{T}_e} \psi_e(x)^{\lambda-1} dx, \quad \Gamma_{a,e} = \int_{\mathcal{T}_{e,k}} \psi_a(x)^{\eta_e(\lambda-1)} \mathcal{C}(x)^{-(1-\alpha)(\lambda-1)} dx$$

and $\mathcal{C}(x) = (\alpha^\alpha(1-\alpha)^{1-\alpha}q(x)^{1-\alpha})^{-1}$ captures capital costs.

Define labor shares $s_e = \Gamma_e w_e^{-\lambda} Y^{1/\varsigma} y_o^{(\varsigma-1)/\varsigma}/L_e$ and $s_{a,e} = \Gamma_{a,e} w_e^{-\alpha\lambda-(1-\alpha)} Y^{1/\varsigma} y_p^{(\varsigma-1)/\varsigma}/L_e$, where $s_e + s_{a,e} = 1$. Log-differentiating the market-clearing condition yields:

$$d\ln w_e = \frac{d\ln \tilde{Y}_e + s_{a,e} d\ln \Gamma_{a,e} - d\ln L_e}{\lambda s_e + (\lambda\alpha + 1 - \alpha) s_{a,e}} \tag{B.1}$$

where $d\ln \tilde{Y}_e = \frac{1}{\varsigma} d\ln Y + \frac{\varsigma-1}{\varsigma}[s_e d\ln y_o + s_{a,e} d\ln y_p]$ captures scale effects.

To characterize how AI shocks affect wages, we focus on the empirically relevant case:

Assumption 1 (AI Shock)**.** The AI shock improves automation technology for production tasks initially performed by high school graduates with capital, $\mathcal{D}_a \subseteq \mathcal{T}_{H,k}$, where the productivity improvements are sufficiently large that $\psi_a(x) > 1$ for all $x \in \mathcal{D}_a$. Additionally, graduates allocate a larger share of their labor to production than dropouts: $s_{a,H} > s_{a,D}$.

Proof of Proposition 1 (Wage Compression). Since $w_D \leq w_H$, we have $\hat{\psi}_a^D < 1$. Assumption 1 requires $\psi_a(x) > 1$ for all $x \in \mathcal{D}_a$, which exceeds the threshold for displacing graduates. Therefore, all affected tasks are reallocated: either to dropouts with capital ($\mathcal{D}_{D,k} \cap \mathcal{T}_{H,k} \neq \emptyset$) or to full automation ($\mathcal{D}_k \cap \mathcal{T}_{H,k} \neq \emptyset$). Since the shock affects only tasks initially performed by graduates, $\mathcal{D}_a \subseteq \mathcal{T}_{H,k}$, no dropout tasks are automated: $\mathcal{D}_k \cap \mathcal{T}_{D,k} = \emptyset$.

The AI shock affects $\Gamma_{a,e}$ through task reallocation and capital complementarity. For dropouts:

$$d\ln \Gamma_{a,D} = \underbrace{\frac{\int_{\mathcal{D}_{D,k}\cap\mathcal{T}_{H,k}} \psi_a(x)^{\eta_D(\lambda-1)} \mathcal{C}(x)^{-(1-\alpha)(\lambda-1)} dx}{\Gamma_{a,D}}}_{>0\text{: tasks reallocate from graduates}}$$
$$+ \underbrace{\frac{\eta_D(\lambda-1)}{\Gamma_{a,D}} \int_{\mathcal{D}_{D,k}} \psi_a(x)^{\eta_D(\lambda-1)-1} \mathcal{C}(x)^{-(1-\alpha)(\lambda-1)} d\ln \psi_a(x) dx}_{>0\text{: capital complementarity}}$$

For graduates:

$$d\ln\Gamma_{a,H} = \underbrace{-\frac{\int_{\mathcal{D}_{D,k}\cap\mathcal{T}_{H,k}}\psi_a(x)^{\eta_H(\lambda-1)}\mathcal{C}(x)^{-(1-\alpha)(\lambda-1)}dx}{\Gamma_{a,H}}}_{<0\text{: tasks reallocate to dropouts}} \underbrace{-\frac{\int_{\mathcal{D}_k\cap\mathcal{T}_{H,k}}\psi_a(x)^{\eta_H(\lambda-1)}\mathcal{C}(x)^{-(1-\alpha)(\lambda-1)}dx}{\Gamma_{a,H}}}_{<0\text{: displacement to automation}}$$

Graduate productivity falls strictly ($d\ln\Gamma_{a,H} < 0$) as all affected tasks are displaced, while dropout productivity rises ($d\ln\Gamma_{a,D} > 0$) through task reallocation and capital complementarity. The condition $s_{a,H} > s_{a,D}$ in Assumption 1 guarantees that the positive scale effect from production expansion is larger for dropouts. From equation (B.1), this implies $d\ln w_D - d\ln w_H > 0$. □

B.4 Labor Reallocation

Proof of Proposition 2 (Reallocation Toward Production). Combining task and factor demands yields labor demand for worker type $e \in \{D, H\}$:

$$\ell_e^*(x) = \begin{cases} Y^{1/\varsigma}y_o^{(\varsigma-1)/\varsigma}\psi_e(x)^{\lambda-1}w_e^{-\lambda} & \text{if } x \in \mathcal{T}_e \text{ (office)}, \\ Y^{1/\varsigma}y_p^{(\varsigma-1)/\varsigma}\psi_a(x)^{\eta_e(\lambda-1)}\mathcal{C}(x)^{-(1-\alpha)(\lambda-1)}w_e^{-\alpha\lambda-(1-\alpha)} & \text{if } x \in \mathcal{T}_{e,k} \text{ (production)}. \end{cases}$$

The AI shock operates through two channels. First, under Assumption 1, production tasks shift from graduates to dropouts, expanding $\mathcal{T}_{D,k}$ and therefore increasing the measure of tasks employing dropout labor in production. Second, higher automation productivity raises production output y_p, which appears in the demand for all production workers. Since dropout labor in production rises both extensively (more tasks) and intensively (higher Y_p within each task), while office labor demand is unaffected by the shock, the relative demand for dropout labor in production versus office work increases. □

B.5 Educational Response

Proof of Proposition 3 (Reduced Educational Investment). Proposition 1 establishes that dropout wages rise more than graduate wages: $d \ln w_D > d \ln w_H$. This raises the opportunity cost of schooling, reducing $\Delta V = (w_H - w_D)/\rho - \tau$.

From the logit structure:

$$\frac{\partial \pi_H}{\partial \Delta V} = \frac{\exp(-\Delta V/\sigma)}{\sigma(1+\exp(-\Delta V/\sigma))^2} = \frac{\pi_H(1-\pi_H)}{\sigma}$$

Since $d\Delta V = (dw_H - dw_D)/\rho < 0$, we have:

$$d\pi_H = \frac{\pi_H(1-\pi_H)}{\sigma} \cdot \frac{dw_H - dw_D}{\rho} < 0$$

The magnitude of the response depends on three factors. First, the response is largest when $\pi_H(1-\pi_H)$ is maximized, which occurs at $\pi_H = 0.5$; this is when the most students are marginal between completing and dropping out. Second, impatient students (high ρ) discount future wage gains more heavily, making them more responsive to current labor market opportunities. Third, lower preference heterogeneity (small σ) means students respond more uniformly to changes in economic incentives. □

B.6 Discussion of Modeling Assumptions

On functional forms. The Cobb-Douglas specification within production tasks and CES aggregation across tasks are chosen for tractability. The qualitative results hold more generally. What matters is:

1. AI and skill are substitutes in production: $\frac{\partial^2 \phi}{\partial \psi_a \partial e} < 0$
2. Tasks are imperfect substitutes: $\lambda > 1$
3. Educational choice responds to relative wages

On the key assumption. The inequality $\eta_H < \eta_D$ can be microfounded in several ways. One interpretation is that AI automates the "cognitive" component of machine operation

(the monitoring, diagnosis, and adjustment that require expertise) while leaving the "manual" component to workers. Another is that AI provides real-time guidance (as in industrial copilots) that substitutes for accumulated experience. Experimental evidence from Brynjolfsson et al. (2025), Noy and Zhang (2023), and Dell'Acqua et al. (2023) supports this pattern across diverse settings.