Entrepreneurial Human Capital and Firm Informality*

Adilya Abdrazakova[a], Luis Franjo[a], and Francesco Turino[a]

[a]Fundamentos del Análisis Económico (FAE), Universidad de Alicante

August 6, 2026

Abstract

This paper studies how entrepreneurial human capital affects firm informality in developing economies. We develop a life-cycle general equilibrium model with endogenous education and occupational choice under limited tax enforcement and credit frictions, in which the labor-market and managerial-ability returns to college education are subject to mismatch risk. Calibrated to Brazil in 2003, the model shows that raising college attainment to the level reached by the country in 2012 reduces informality by reallocating talent toward larger and more productive formal firms, thereby increasing GDP and aggregate productivity. Strikingly, this aggregate decline masks a rise in informal entrepreneurship among the newly skilled themselves, as financial frictions can still make operating informally optimal at early stages of business creation; this residual informality disappears only once credit access improves alongside education. We provide empirical support for the mechanism using microdata, leveraging Brazil's 1996 education reform.

Keywords: Informality, Entrepreneurship, Human Capital, Financial Frictions, Misallocation, Tax Evasion.

JEL Classification: E26, E44, H26, I25, J24, L26, O17.

*We would like to thank M. Dolores Collado, Omar Rachedi, Andrés Erosa, Henry Redondo, Alexander Ludwig, Anna Sanz de Galdeano, Pedro Albarran, Lian Allub, Sang Yoon (Tim) Lee, Teresa Molina-Millán, Antonia Díaz, Alexander Monge-Naranjo, Francesco Serti, Hardi Ahmed, Lluc Puig, Selene Perazzini, Julian Messina, Antonio Minniti, Davide Dragone, Nicola Pavoni, Markus Poschke and seminar participants at University of Western Australia, University of Bologna, Queen Mary University of London, University of Southampton, World Bank Workshop on Informality in the Latin America and the Caribbean Region, SIE Annual Meeting, CREI@30 conference, Inequalities in Low- and Middle-Income Countries workshop at EUI, VII Workshop of the Spanish Macroeconomics Network, UC3M Economics PhD 35th Anniversary Alumni Conference, Alicante Workshop in "Macroeconomics and Productivity", Lucca Macro Workshop, and International Workshop on Economic Growth and Macroeconomic Dynamics. Luis Franjo acknowledges financial support from the Catedra Germán Bernácer. Luis Franjo and Francesco Turino acknowledge financial support from Consellería de Innovación, Universidades, Ciencia y Sociedad Digital de la Generalitat Valenciana (CIPROM/2021/060). Adilya Abdrazakova also acknowledges financial support from Consellería de Innovación, Universidades, Ciencia y Sociedad Digital de la Generalitat Valenciana (CIGRIS/2021/013).

1 Introduction

Firm informality is a pervasive feature of low- and middle-income economies (La Porta and Shleifer, 2014). In Brazil, for example, available estimates indicate that approximately two-thirds of firms operate without registration with tax authorities, absorbing nearly 35% of urban employment and generating output equivalent to between 30% and 40% of official GDP (Ulyssea, 2018; Medina and Schneider, 2018). A large body of literature has proposed various explanations for the persistent dominance of informal firms in developing countries, including weak regulatory enforcement, high entry costs, financial frictions, and the abundance of unskilled labor (Amaral and Quintin, 2006; Antunes and Cavalcanti, 2007; Haanwinckel and Soares, 2021; Erosa *et al.*, 2023).

One potential factor that has received comparatively less attention is the role of entrepreneurial human capital. Compared to formal firms, informal enterprises tend to be smaller, more stagnant, and are typically managed by individuals with lower levels of education (de Paula and Scheinkman, 2007; La Porta and Shleifer, 2008; Ulyssea, 2018). These empirical regularities—combined with new microdata evidence highlighting the importance of entrepreneurs' education for firm dynamics (Queiró, 2021)—raise the question of whether insufficient entrepreneurial human capital contributes to firm informality in developing countries, where large informal sectors and low educational attainment often coexist (La Porta and Shleifer, 2014). If so, what are the implications of this firm-level linkage for aggregate output and productivity?

To answer these questions, this paper develops a quantitative life-cycle general equilibrium model of occupational choice with endogenous college enrollment decisions, limited tax enforcement, and imperfect credit markets.[1] In this framework, educational choice plays a central role in shaping occupational outcomes, as acquiring a college degree affects individuals' skills through two distinct channels. First, consistent with traditional models of human capital accumulation (Heathcote *et al.*, 2017), college-educated individuals allocate their time endowment to the skilled labor market. Second, as a relatively novel feature in the literature (Berniell, 2021; Queiró, 2021; Morazzoni, 2021), we model entrepreneurial human capital by assuming that college education enhances managerial ability. Neither benefit is guaranteed, however: with an exogenous probability, a college graduate experiences educational mismatch and fails to obtain the skilled wage and the managerial-ability gain typically associated with a degree. We show that, in this environment, entrepreneurial human capital has quantitatively important implications for GDP, TFP, and the extent of firm informality, with the magnitude of these effects shaped by the severity of financial frictions.

Our model is similar in spirit to Buera (2009), Buera and Shin (2013), and Franjo *et al.* (2022), featuring financially constrained, heterogeneous individuals who decide in each period whether to work as paid employees or operate an entrepreneurial activity, either formally or informally. Formal entrepreneurs comply with business regulations and pay taxes in exchange for access to external credit. Informal entrepreneurs, by contrast, are non-compliant individuals who evade taxes and rely exclusively on internal funds, while facing a probability

[1] Evidence from the World Bank Enterprise Surveys, as reported in La Porta and Shleifer (2014), shows that access to finance is perceived as the most significant obstacle to doing business by both formal and informal entrepreneurs in low- and middle-income countries.

of detection by fiscal authorities that rises with firm size. We enrich this framework by incorporating capital-skill complementarity in production (Krusell *et al.*, 2000) and endogenous educational choice, the latter modeled as a pre-working decision in which each individual determines whether to invest in a costly college degree or maintain a lower level of schooling.

We calibrate the model to match key features of the Brazilian economy in 2003, including the size of the informal sector, college completion rates, earnings-skill premia for both workers and entrepreneurs, share of college graduates employed outside high-skill occupations, and the firm size distribution for both formal and informal enterprises. The structure of the calibrated model is sufficiently rich to generate substantial heterogeneity in occupational choices conditional on education, and to rationalize a prominent feature of the data: informal entrepreneurship is predominantly concentrated among less-educated individuals.

The characterization of the steady-state equilibrium reveals that this heterogeneity is primarily driven by entrepreneurial human capital. Although agents in the model have access to the same production technology, college-educated entrepreneurs who avoid mismatch have enhanced managerial ability, so—consistent with the empirical evidence in Queiró (2021)—their firms enter the market at a larger size and grow faster over the life cycle than those run by equally talented but non-college-educated entrepreneurs. Since the probability of detection increases with firm size, this translates into lower expected returns from tax evasion and a stronger incentive to operate formally. This effect is reinforced by the wage-skill premium, which raises the opportunity cost of entrepreneurship for matched college-educated individuals and induces most of them to work as skilled employees rather than start small, informal firms. As a result, informality becomes concentrated among less-educated entrepreneurs, who operate small, financially constrained firms and face relatively low expected penalties for non-compliance.

To assess the aggregate consequences of these channels, we conduct a counterfactual experiment that brings the college attainment rate in the calibrated economy up to the level Brazil actually reached in 2012, after a decade of rapid educational expansion following the 1996 Higher Education Reform. This realistic increase in college attainment—from 8.4% to 12.5% of the population—generates sizable long-run gains for the Brazilian economy: official GDP rises by 7.7%, measured TFP by 2.3%, and fiscal revenues by 8.6%. Most importantly, the size of the informal sector declines by 11.3 percentage points, from 34.5% to 23.2% of official GDP.

This decline is driven by a shift in the skill composition of the labor force, which lowers the wage-skill premium and simultaneously encourages paid employment among non-college individuals and entrepreneurship among the newly college-educated, who benefit from enhanced managerial ability. We show that this latter channel—entrepreneurial human capital—is central to these results: in a counterfactual economy where college education raises only wages, and not managerial ability, the gains in official GDP and measured TFP are only about half as large.

Perhaps counterintuitively, we show that this aggregate decline in the size of the informal sector masks a rise in informality among the newly skilled themselves: even though matched college-educated entrepreneurs run larger, more productive firms, financial frictions can still make informal entrepreneurship the optimal way to finance a firm's early years, so the mass of skilled informal entrepreneurs increases even as aggregate informal production shrinks. A robustness exercise that pushes college attainment among entrepreneurs to near-universal levels confirms this mechanism: firm informality does not vanish, but instead becomes increasingly concentrated among skilled entrepreneurs as the cost of college falls. This residual informality closes only when credit access improves alongside educational attainment: as financial frictions are relaxed, informal entrepreneurship among the skilled falls toward zero regardless of the college attainment rate, and the aggregate gains from higher educational attainment are correspondingly amplified. The policy implication is clear: fully exploiting the formalization gains from education in developing economies requires addressing financial and educational barriers jointly, not in isolation.

Finally, to empirically test the main predictions of the model, we draw on Brazil's 1996 Higher Education Reform, which reduced barriers to entry for for-profit higher education institutions (Cox, 2024). Following the empirical strategy of Duflo (2001), we exploit variation in exposure to the reform across cohorts and commuting zones, using excess college entry as a proxy for treatment intensity in a difference-in-differences framework. We find that individuals in areas with greater excess college entry were significantly less likely to engage in informal entrepreneurship. Altogether, these results support the model's prediction that expanding access to higher education is a powerful lever for shifting occupational choices away from informality.

Related literature Our paper contributes to the literature on economic development, financial frictions, and informality by incorporating endogenous human capital accumulation into a general equilibrium model with occupational choice. We build on a well-established tradition of research that emphasizes the importance of financial frictions for development, including Erosa (2001), Buera (2009), Buera *et al.* (2011), Buera and Shin (2013), and Allub and Erosa (2019). These studies show that limited access to credit leads to misallocation of capital, thereby reducing income per capita and aggregate productivity. Extending this line of work, a parallel strand of the literature introduces informality into models with financial frictions to better capture the realities of developing economies. Papers such as Amaral and Quintin (2006), Antunes and Cavalcanti (2007), Quintin (2008), D'Erasmo and Boedo (2012), D'Erasmo (2013), Franjo *et al.* (2022), and Erosa *et al.* (2023) incorporate informal production as a response to limited financial access and enforcement capacity. Our paper builds directly on this line of work by incorporating endogenous human capital into the analysis. This extension is important because it highlights that the capacity of entrepreneurs to benefit from formalization depends not only on institutional and financial constraints but also on their ability to manage firms efficiently, a margin that is endogenously determined and shaped by education decisions.

While prior research has studied the interaction between human capital and informality, the focus has largely been on workers' education within search and matching frameworks, as in Haanwinckel and Soares (2021), Bobba *et al.* (2021), and Bobba *et al.* (2022). We complement and extend this literature by shifting the focus to entrepreneurial human capital and showing how this margin plays a first-order role in shaping informality and development outcomes. The paper most closely related to ours is Berniell (2021), who analyzes theoretically the decision of entrepreneurs to acquire education in the presence of financial constraints. While our model is quantitative in nature, we also explore the implications of entrepreneurial human capital for informality and aggregate outcomes. In particular, our model quantifies how the limited accumulation of entrepreneurial human capital constrains firm growth and the formalization process. This contribution is essential because it shows that low levels of managerial ability, driven by constrained educational investment among potential entrepreneurs, are a central obstacle to formalization, offering a new explanation for persistent informality beyond labor market frictions and enforcement limitations.

Finally, our paper also complements recent work such as Allub *et al.* (2023), which explores capital-skill complementarity and workers' education in developing economies with financial frictions. We show that the interaction between financial access and entrepreneurial human capital is equally critical: formalization and productivity gains depend not only on access to capital and skilled labor, but also on who becomes an entrepreneur and the quality of their managerial capacity. This distinction is important because it highlights a different channel through which education and financial development interact: not only by increasing the productivity of workers, but also by enabling more talented and better-educated individuals to start and expand formal firms, thus amplifying aggregate gains from human capital accumulation and credit market reforms. In this way, our paper contributes a novel mechanism to the broader literature on informality, development, and human capital accumulation.

Outline The rest of the paper is organized as follows. Section 2 presents key stylized facts, while Section 3 describes the theoretical framework. Section 4 outlines the calibration strategy and evaluates the model's fit for both targeted and non-targeted moments. Section 5 derives the main firm-level mechanisms and assesses aggregate implications through a series of counterfactual experiments. Section 6 provides empirical evidence supporting the model's main predictions, and Section 7 concludes. Additional details on the data, calibration, and further model results are provided in the Online Technical Appendix.

2 Background

In this section, we present key facts about the relationship between informality and education. While we primarily focus on the Brazilian economy, which serves as the benchmark for our analysis, we also investigate this relationship from a cross-country perspective. We begin by introducing precise definitions for the main variables of interest, then present a combination of micro and macro data, and finally provide statistics on informality and education that characterize the main stylized facts. These facts provide empirical support for our modeling assumptions in Section 3 and inform our quantitative exercises in Sections 4 and 5.

2.1 Definitions and data description

Throughout the paper, we define the *informal economy* as all production activities of a country involving legal goods and services that are deliberately concealed from fiscal authorities to evade taxes. In line with Medina and Schneider (2018), this definition does not include illegal economic transactions−such as drug trafficking−or home production, thereby establishing a one-to-one relationship between informality and tax evasion. Consistently, we classify *informal firms* as enterprises that are not registered with tax authorities and *informal workers* as employees who do not hold a regular labor contract. Finally, we define *college-educated workers* and *college-educated entrepreneurs* as employees and managers, respectively, who have completed tertiary education.

We use two main datasets specific to the Brazilian economy: the ECINF (*Pesquisa de Economia Informal Urbana*) and the PNAD (*Pesquisa Nacional por Amostra de Domicílios*) surveys. The former is a repeated cross-sectional survey of small businesses conducted by the Brazilian Bureau of Statistics (IBGE) in 1997 and 2003. The main advantage of ECINF is that it is a matched employer-employee database, providing detailed information on both firm and worker informality. Specifically, firms are asked whether they hold the tax identification number required for Brazilian businesses (i.e., registration in the *Cadastro Nacional de Pessoa Jurídica*) and whether each of their employees holds a formal labor contract.[2] Thus, the ECINF data set allows the identification of informal production units in the sample, providing valuable information on the characteristics of their managers and workers.

However, since the main objective of the ECINF survey was to measure the role and operational characteristics of informal activity in Brazil, IBGE applied sampling procedures designed to ensure national representativeness for small firms (up to five employees) and informal businesses. To complement this, we use the PNAD, a nationally representative household survey that provides detailed information on labor market outcomes (e.g. earnings, employment), sociodemographic characteristics (e.g. age, level of education), and informality status. For comparability, we restricted our sample to 2003, as it is the last available year of ECINF. One drawback of PNAD 2003 is that it does not distinguish between formal and informal entrepreneurs, a limitation that was addressed starting in 2012 with the introduction of PNAD *Contínua* (PNAD-C). Consequently, in this section we rely on PNAD 2003 to primarily characterize labor informality, while using the 2012 release when additional insights are needed. Further details on the ECINF and PNAD datasets can be found in Section C of the Online Technical Appendix.

We also conduct a cross-country analysis by merging two databases: the Global Entrepreneurship Monitor (GEM) dataset and the Medina and Schneider (2018) estimates on the size of the informal economy around the world. The GEM's objective is to measure the level and nature of entrepreneurial activity worldwide by collecting nationally representative samples of entrepreneurs across countries. We restrict our sample to the years 2009–2015 to leverage a harmonized variable on education levels and to focus on non-agricultural, non-nascent entrepreneurs.[3] Additionally, we include only countries with at least three years of available observations, resulting in a sample of 54 countries. This database is then merged with the Medina and Schneider (2018) dataset, which provides yearly estimates of the size of the informal sector in 158 countries over the 1991–2015 period. The methodology used to estimate these figures is based on a modified version of the standard Multiple Causes, Multiple Indicators approach.

2.2 Stylized Facts

Fact 1: The informal sector is more intensive in less-educated workers and entrepreneurs. There is substantial evidence in the literature that informal firms, on average, are run by

[2]In Brazil, firms are required to record a worker's employment history in an employment booklet known as a *carteira de trabalho*. A salaried position is considered formal if it is registered in this booklet.

[3]GEM defines entrepreneurship as "Any attempt at new business or new venture creation, such as self-employment, a new business organization, or the expansion of an existing business, by an individual, a team of individuals, or an established business."

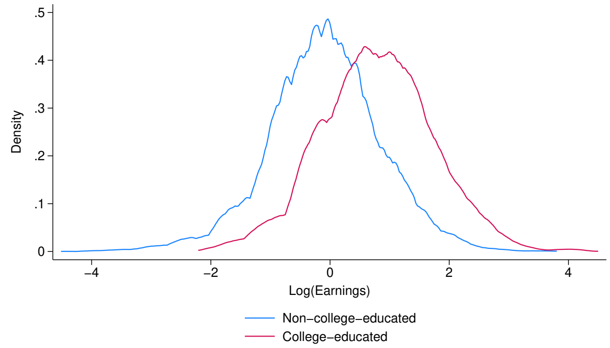
Table 1: Informality and education

Moments	Source	Value
Panel A: Educational attainment of working aged population		
(A.1) *College rate*:		
All individuals	PNAD 2003	8.36
Among entrepreneurs	PNAD 2003	10.16
Among workers	PNAD 2003	7.64
Panel B: Small businesses		
(B.1) *Share of college-educated entrepreneurs*:		
In formal firms	ECINF 2003	17.45
In informal firms	ECINF 2003	6.28
(B.2) *Informality rate*:		
Among college-educated entrepreneurs	ECINF 2003	57.13
Among non-college-educated entrepreneurs	ECINF 2003	80.78
Panel C: All firms		
(C.1) *Share of college-educated workers*:		
In formal firms	PNAD-C 2012	13.16
In informal firms	PNAD-C 2012	2.88
(C.2) *Informality rate*:		
Among all workers	PNAD 2003	26.91
Among college-educated workers	PNAD 2003	17.14
Among non-college-educated workers	PNAD 2003	27.72
(C.3) *Share of college-educated entrepreneurs by firm size*:		
1-5 workers	PNAD 2003	18.86
6-10 workers	PNAD 2003	28.46
$>$ 10 workers	PNAD 2003	36.77

less-educated entrepreneurs and tend to hire more unskilled workers compared to formal firms (La Porta and Shleifer, 2014). These features are also apparent in Brazilian data. As shown in Panel B.1 of Table 1, informal firms in the ECINF dataset have a lower share of college-educated entrepreneurs relative to formal firms of comparable size. This trend is further corroborated by Panel B.2, which shows that informality is more prevalent among non-college-educated entrepreneurs. Specifically, among small businesses, 80% of entrepreneurs with low educational attainment were running informal firms in Brazil in 2003. This figure is about 24 percentage points higher than the corresponding statistic for college-educated entrepreneurs. Similarly, Panel C.2 shows that the informality rate is generally higher among non-college-educated workers, confirming that the informal sector is more intensive in unskilled labor as shown in Panel C.1. This evidence, coupled with the fact that (i) informal firms tend to be smaller in size compared to formal firms (de Paula and Scheinkman, 2007; La Porta and Shleifer, 2008; Leal Ordoñez, 2014; Ulyssea, 2018), and (ii) formal and informal firms coexist within narrowly defined industries (Ulyssea, 2018), is consistent−among other possible explanations−with a theoretical model in which formal and informal enterprises have access to an identical production technology featuring capital-skill complementarity. Further support for this explanation comes from Fonseca and Van Doornik (2022), who provide evidence that capital-skill complementarity is an empirically relevant feature of production technologies in Brazil.

Fact 2: Educational attainment of entrepreneurs is positively correlated with firm size. Extending the previous analysis, we now assess how the educational attainment of entrepreneurs correlates with firm size. Due to the size cap of the ECINF survey, we address this question using the PNAD dataset, which includes medium and large firms. We begin by reporting college attainment rates among entrepreneurs by firm size. We compute these statistics for firms in general (i.e., without distinguishing for informality status), using the number of paid workers as a measure for firm size. As illustrated by Panel C.3 of Table 1, we observe a clear upward trend in the share of college-educated entrepreneurs as firm size grows. While only

Figure 1: **Firm size distributions**



Notes: Data are taken from PNAD 2003. To compute the reported distributions, we first regress log-earnings, normalized by the average across working-age population, on industry dummies to control for inter-industry variability. We then use the residuals from this regression to compute the weighted kernel density of log-earnings by college attainment.

19% of entrepreneurs running small firms in Brazil were college educated in 2003, this figure rises to 28% for medium-sized firms and nearly doubles for large firms. Additional evidence of this firm heterogeneity is presented in Figure 1, which shows that the distribution of earnings for college-educated entrepreneurs—as a proxy for firm size—is right-shifted relative to that for non-college individuals. In other words, our findings indicate that, on average, firms run by college-educated entrepreneurs tend to be larger. This evidence is consistent with the results provided by Queiró (2021), who, through analysis of Portuguese administrative data, also documents a positive relationship between educational attainment of entrepreneurs and firm size—both in life-cycle behavior and in cross-sectional firm heterogeneity. Additionally, Queiró investigates the mechanisms underlying this empirical pattern, providing evidence that technology adoption and innovation may be the primary drivers linking firm size, and productivity, with entrepreneurial human capital accumulation through schooling. Similar conclusions are drawn in Ciccone and Papaioannou (2009).

Fact 3: College and formal earning premia are substantial. As shown in Panel A of Table 2, in 2003 there were substantial earnings gains from completing tertiary education in Brazil. On average, a college-educated entrepreneur and a college-educated worker earned approximately 3.9 and 3.5 times more, respectively, than their non-college-educated counterparts. Similarly, we find significant earnings premia for individuals operating in the formal sector, with entrepreneurs and workers earning, on average, 2.44 and 1.70 times more than informal employers and employees, respectively. All of the reported statistics are in line with the evidence provided by de Paula and Scheinkman (2011), who, using ECINF 2003 data in a *Mincerian* regression, found positive and statistically significant returns from both formalization and education for Brazilian entrepreneurs. Our results further highlight that, on average, skill and formal premia for entrepreneurs are larger relative to those for workers. These findings are consistent with the evidence reported by Berniell (2021), who documents that education returns for entrepreneurs tend to be higher compared to workers in countries with large informal sectors.

Fact 4: Across countries, the size of the informal sector is negatively correlated with the share

Table 2: Earnings gaps

Moments	Source	Value
(A) *College earning premium*:		
Entrepreneurs	PNAD 2003	3.93
Workers	PNAD 2003	3.55
(B) *Formal-informal earning gap*:		
Entrepreneurs	ECINF 2003	2.44
Workers	PNAD 2003	1.70

Note: College and formal earning premia are computed as ratio of average earnings.

of college-educated entrepreneurs. Another interesting feature of the Brazilian economy, apparent in our data, is its combination of low educational attainment among entrepreneurs and high informality rates. In 2003, only 10.16% of all entrepreneurs were college-educated (Panel A of Table 1),[4] and around 66% of all firms operated in the informal sector (Ulyssea, 2018), with nearly 27% of all salaried employees not holding a regular labor contract (Panel C.2 of Table 1).[5] Using the GEM and Medina and Schneider (2018) databases, we find that this characteristic holds true across countries, with the share of college-educated entrepreneurs being negatively correlated with the size of the informal economy. This finding is clearly visible in Figure 2, which shows that, on average, countries with a larger share of output produced informally tend to have lower college rates among entrepreneurs.

3 Model

We consider an overlapping generations model of entrepreneurship with endogenous educational choices, taxation and financial market frictions.[6] Each generation consists of individuals who are heterogeneous in terms of innate talent, costs of education and wealth profiles, the latter being endogenously determined by forward-looking saving decisions. We also allow for educational mismatch, occurring with probability $\chi \in [0, 1)$, whereby college education fails to yield its associated earnings premium.[7]

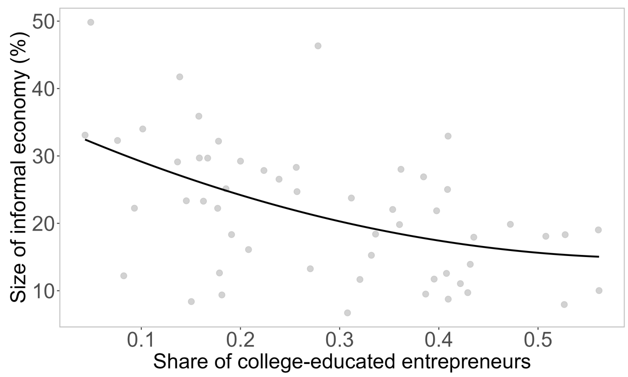
[4]This number reflects the low college attainment rate in the Brazilian population, where only 8.36% of working-aged individuals were college-educated in 2003. It is also interesting to note that the college rate among entrepreneurs was 1.32 times larger compared to that of workers (see Panel A of Table 1). This feature is consistent with the findings of Berniell (2021), who documents that in countries with relatively large informal sectors, a larger fraction of skilled individuals choose to become entrepreneurs.

[5]According to the Medina and Schneider (2018), these numbers translate into a total output produced by informal firms equivalent to a 39% of the official Brazilian GDP in 2003.

[6]The overlapping generations structure of the model is motivated by compelling empirical evidence highlighting the importance of the life cycle stage of firms and entrepreneurs in influencing informality rates among enterprises in developing countries (Perry *et al.*, 2007; de Paula and Scheinkman, 2010, 2011; Diaz *et al.*, 2018).

[7]Educational mismatch, whereby college graduates do not obtain occupations or earnings typically associated with a college degree, is a well-documented feature of Brazilian and Latin American labor markets. Using matched employer-employee data, Marioni (2021) finds that around 25% of Brazilian workers are mismatched relative to their educational attainment, with overeducated workers earning significantly less than comparable well-matched co-workers. Similar evidence exists for other Latin American countries that, like Brazil, experienced rapid expansions in college enrollment (Castro *et al.*, 2024). Motivated by this evidence, we allow for the possibility that a college degree does not translate into the labor-market returns associated with college education.

Figure 2: **Entrepreneurial college rates and size of the informal economy**



Notes: The grey dots represent average values for cross-country data over the 2009–2015 period. Fitted values from a quadratic regression are reported with a continuous black line.

The life cycle of each individual is characterized by three distinct stages: education, work and retirement. The first stage, labeled age 0, is a pre-working period of life in which the individual chooses the education level $h = \{s, u\}$, where s denotes college education and u non-college education. In the subsequent work stage, an individual decides in each period whether to become a worker or to run an entrepreneurial activity, on the basis of her financial wealth, her occupational skills, and her education attainment. Workers supply their time-endowment inelastically and receive a gross wage conditional to their labor-market skill type while entrepreneurs operate a technology that combines their managerial abilities with skilled labor, unskilled labor, and capital to produce output. Individuals enter the labor market at age 1, retire at the mandatory age J_R and then die when they reach the age $J \in (J_R, \infty)$.

As in Buera and Shin (2013), imperfections in financial markets manifest as collateral requirements on capital rentals. Tax enforcement is also imperfect, as government can detect informal economic activities−those carried out by unregistered firms−only upon inspection. Entrepreneurs can thus avoid taxation by operating informally. However, non-compliance with tax payments incurs costs. Firstly, access to credit requires registration with tax authorities, meaning that informal entrepreneurs must operate in a financial autarky regime. Secondly, the government conducts random audits on a number of individuals each period, so informal entrepreneurs face a probability of being detected and forced to pay the taxes due, augmented by a penalty surcharge.

3.1 Time convention, preferences and endowments

Time is discrete, and at the beginning of every period, a new generation of individuals is born. Agents discount the future exponentially using a common discount factor $\beta \in (0, 1)$, and have instantaneous preferences over consumption that are represented by a CRRA utility function of the form

$$u(c_j) = \frac{c_j^{1-\sigma}}{1-\sigma}$$

where c_j denotes consumption at age $j = \{1, 2, ..., J\}$, and $\sigma > 0$ is the coefficient of relative risk aversion.

A newly born individual is endowed with (i) zero assets; (ii) an innate talent $e_u \in \Theta$, drawn from an invariant distribution with CDF $\Phi(e_u)$, which determines the managerial productivity of non-college-educated entrepreneurs, and (iii) one unit of time that can be supplied in the market for unskilled labor in each period of the working stage. As described in a separate section below, labor and managerial skills remain unchanged throughout the life cycle unless the agent acquires a college degree during the educational stage, which, with probability $1-\chi$, affects innate abilities in two ways. First, college-educated individuals can allocate their time endowments to the skilled labor market. Second, higher education enhances managerial productivity with an impact that is individual-specific. To distinguish between the productivity of college- and non-college-educated entrepreneurs, in what follows, we refer to e_h as the managerial ability of an individual with education level $h = \{s, u\}$.

3.2 Firms

Following Quadrini (2000), we assume that there are two distinct sectors producing the same homogeneous consumption good. The first sector, labeled the entrepreneurial sector, is characterized by small-scale enterprises owned by individuals engaged in entrepreneurship. The second one, referred to as the corporate sector, comprises large-scale, impersonal firms. In addition to differences in size, firms run by entrepreneurs and corporations are assumed to differ along two other key dimensions: the possibility of non-compliance with business regulations and the severity of financial constraints.

3.2.1 Entrepreneurial sector

We follow Allub *et al.* (2023) by assuming that an entrepreneur combines her managerial ability, e_h, with capital, k, skilled labor, ls, and unskilled labor, lu, to produce output via the following CES technology

$$f(e_h, lu, ls, k) = e_h^{\eta} \left[\mu lu^{\theta} + (1-\mu)(\iota k^{\rho} + (1-\iota) ls^{\rho})^{\frac{\theta}{\rho}} \right]^{\frac{1-\eta}{\theta}} \tag{1}$$

where μ, ι, $\eta \in (0,1)$ are distribution parameters, while θ, $\rho \leq 1$ control for the elasticities of substitution between unskilled labor, capital, and skilled labor. This production function is a variant of the technology proposed by Krusell *et al.* (2000) and exhibits decreasing returns to scale, thereby capturing the notion of *span of managerial control* popularized by Lucas (1978). Additionally, provided that $\theta > \rho$, it also embodies capital-skill complementarity.[8]

Unskilled and skilled workers are hired from perfectly competitive labour markets at the wage rates w_u and w_s, respectively, while capital is financed by entrepreneurs using their own assets, a, and, in addition, external resources borrowed from financial intermediaries. We assume that to access credit, entrepreneurs need to register with tax authorities, which makes their production activities observable by the government. Entrepreneurs who comply with these regulations are labeled formal. In contrast, informal entrepreneurs are non-compliant individuals who consequently choose to operate their businesses without resorting to external financing. These entrepreneurs, however, may successfully evade taxation by hiding their

[8]More specifically, under the parameterization $\theta > \rho$, the production function features higher elasticity of substitution between capital and unskilled labor (i.e., $1/(1-\theta)$) than between capital and skilled labor (i.e., $1/(1-\rho)$).

production activities, which can only be detected by the government after a monitoring process.

An implicit assumption of the production function specified in Equation (1) is that, apart from the entrepreneur's managerial ability, there is no difference in the technology adopted by formal and informal firms, with both types demanding capital, skilled labor, and unskilled labor. In other words, firms are ex-ante identical, and any difference in productivity, size and skills intensity between formal and informal firms endogenously arises in equilibrium from the self-selection of talent into entrepreneurship.[9] To distinguish between formal and informal factor demands, in what follows we use the subscript f for variables associated with formal entrepreneurs and i for variables associated with informal entrepreneurs.

3.2.2 Corporate sector

Following Franjo *et al.* (2022), we assume that in the corporate sector operate a large number of perfectly competitive firms. These enterprises are heterogeneous in productivity and use the same production technology as firms in the entrepreneurial sector. In addition, we assume that firms in the corporate sector (i) pay an operational fixed cost; (ii) cannot engage in informal activities, and (iii) are not subject to collateral constraints on capital rentals.[10,11] In the Online Technical Appendix A, we show that under these assumptions the aggregate net output of the sector, Y_c, is given by

$$Y_c = A\left[\mu Lu_c^{\theta} + (1-\mu)(\iota K_c^{\rho} + (1-\iota)Ls_c^{\rho})^{\frac{\theta}{\rho}}\right]^{\frac{1-\eta}{\theta}} - \phi_f \tag{2}$$

where $A > 0$ is a parameter controlling for the sector-specific productivity; K_c, Lu_c and Ls_c respectively stand for aggregate capital, aggregate unskilled labor and aggregate skilled labor in the corporate sector, while $\phi_f > 0$ is the operational fixed cost aggregated across firms.

3.3 Financial intermediaries

We assume that there is a continuum of perfectly competitive financial intermediaries, each receiving deposits from households at a risk-free interest rate, r, and renting capital to firms at a rental rate, r_k. In equilibrium, the zero-profit condition implies

$$r_k - \delta = r \tag{3}$$

[9] In addition to the stylized facts documented in Section 2, this implicit assumption is also supported by evidence that challenges the dualistic view of formal and informal firms. See for example Meghir *et al.* (2015), Allen *et al.* (2018), Ulyssea (2018) or Haanwinckel and Soares (2021).

[10] In our model, financial frictions and firm informality depress capital demand, potentially preventing the clearing of capital market. By incorporating large corporations that are not subject to borrowing constraints, we introduce in the model an additional source of capital absorption that resolves this issue and ensures the existence of general equilibrium for any model parameterization. For a more detailed discussion on this topic, interested readers are referred to Franjo *et al.* (2022).

[11] The assumption on the operational fixed cost has been introduced to guarantee that corporate profits are zero in equilibrium.

where $\delta \in (0,1)$ is the capital depreciation rate. Following Buera and Shin (2013), we assume that there is limited contract enforceability for firms in the entrepreneurial sector so that entrepreneurs' capital rental is limited by a collateral constraint of the form

$$k \leq \lambda a \tag{4}$$

where a stands for the individual financial wealth, and $\lambda \geq 1$ measures the degree of financial imperfections. In particular, $\lambda = \infty$ corresponds to perfect capital markets, while $\lambda = 1$ denotes financial autarky. The latter case describes the financing regime of an informal entrepreneur, who has no access to credit and therefore can only self-finance capital with her accumulated wealth, i.e.,

$$k_i \leq a \tag{5}$$

3.4 Government

The government levies flat taxes τ_c and τ_y on consumption and personal income to finance wasteful public expenditures. Additionally, to discourage tax evasion, the government carries out random audits of individuals and requires detected informal entrepreneurs to pay the taxes due, augmented by a penalty surcharge factor $s > 0$. We assume that the probability of being audited in a given period, denoted by $p(k_i)$, increases with the amount of capital used by an informal firm. This assumption is common in the literature on informality and tax evasion (Leal Ordoñez, 2014; Di Nola *et al.*, 2021; Franjo *et al.*, 2022; Fernández-Bastidas, 2023) and can be rationalized by the fact that large establishments are more visible to tax authorities, making it harder for them to hide production compared to smaller ones.

3.5 Life cycle

3.5.1 Education

At the beginning of her life (i.e., at age 0), an individual makes a discrete choice regarding human capital investment, choosing between pursuing a college degree ($h = s$) or maintaining a lower level of schooling ($h = u$). College education may enhance occupational skills along two dimensions: access to the skilled labor market and higher managerial ability. These benefits, however, are subject to mismatch risk. At the start of the working stage—after the education cost has been incurred—each college-educated individual draws a binary mismatch indicator $m \in \{0,1\}$, with $\Pr(m = 1) = \chi \in [0,1)$.

If the individual is *matched* ($m = 0$), she obtains the full returns from college. First, she is recognized, and paid, as a highly skilled worker in the labor market. Second, her managerial skills improve according to the following technology

$$e_s = e_u + \psi e_u^{\epsilon} \tag{6}$$

where $\epsilon > 0$ controls for the curvature of the ability-improvement function, and $\psi > 0$ determines the extent to which college education translates into human capital accumulation.[12]

[12]The assumption that the enhanced skill e_s positively depends on the innate talent e_u captures the notion of self-productive investment in skill formation popularized by Cunha and Heckman (2007), and is consistent with the empirical observation that talented individuals may be more effective in accumulating

If the individual is *mismatched* ($m = 1$), college does not translate into higher occupational skills. Her effective managerial ability remains e_u, and she can only supply labor in the unskilled market, earning wage w_u. In terms of labor market outcomes, a mismatched college-educated individual is therefore economically equivalent during the working stage to a non-college-educated individual. We nonetheless keep educational attainment, h, distinct from the mismatch realization: a mismatched individual is still recorded as college-educated ($h = s$) when computing model statistics, since she completed college but failed to convert it into higher labor-market returns.[13]

We also assume that each individual faces an idiosyncratic utility cost of attending college, given by

$$\zeta\left(e_u, \kappa\right) = \kappa e_u^{-\phi} \tag{7}$$

where $\phi > 0$, and κ denotes a stochastic component drawn from a time-invariant distribution with CDF $\Gamma(\kappa)$. This specification extends the formulation proposed by Heathcote *et al.* (2010) by assuming that the educational cost decreases with the individual's innate ability e_u.[14] Regarding the stochastic component κ, we instead maintain the original interpretation of Heathcote *et al.* (2010) by viewing the distribution $\Gamma(\kappa)$ as accounting, in a reduced form, for psychological and pecuniary costs of education (e.g. variations in parental income, government aid programs, tuition fees).

Under the above assumptions, an individual decides whether to invest in human capital by considering three factors: (i) her draw for the stochastic component of the educational cost, κ; (ii) her innate talent, e_u; and (iii) the expected college earnings premia. More formally, let $\mathbb{V}^{e_u}(s)$ denote the expected lifetime utility of a *matched* college-educated individual with innate ability e_u, and $\mathbb{V}^{e_u}(u)$ denote the expected lifetime utility of a non-college-educated individual with the same innate ability. Since a mismatched individual is economically equivalent to a non-college-educated individual in the working stage, her continuation value equals $\mathbb{V}^{e_u}(u)$. The expected continuation value from attending college is therefore $(1-\chi)\mathbb{V}^{e_u}(s) + \chi\mathbb{V}^{e_u}(u)$. The optimal schooling decision for an individual with stochastic educational cost κ is then:

$$h(e_u, \kappa) = \begin{cases} s & \text{if } (1-\chi)\mathbb{V}^{e_u}(s) + \chi\mathbb{V}^{e_u}(u) - \zeta\left(e_u, \kappa\right) \geq \mathbb{V}^{e_u}(u) \\ u & \text{otherwise} \end{cases}$$

Equivalently, an individual attends college if and only if the expected return from a successful

human capital through schooling (e.g. Belzil and Hansen, 2002).

[13]This distinction is what separates a mismatched college-educated individual from a non-college-educated individual ($h = u$) with the same realized ability and wage, the latter never having incurred the cost of college.

[14]The assumption that the cost of acquiring a college degree depends on e_u introduces additional heterogeneity into the model and, to some extent, links a worker's productivity to her innate talent. To see this, consider two different individuals who both decide to remain workers throughout their entire life cycle and have drawn the same κ. Since the cost of attending college decreases with e_u, these individuals, despite not benefiting from the improvement in managerial ability, may still choose different education levels due to their distinct innate talents and, as a result, earn different wages. In other words, the assumed specification for the cost of education establishes a connection between the formation of a worker's skills through schooling and her ability endowment e_u. This property, coupled with the ability-improvement technology (6), allows us to interpret more broadly the stochastic variable e_u as innate talent rather than merely as managerial ability.

match exceeds the cost of education:

$$(1-\chi)\left[\mathbb{V}^{e_u}(s)-\mathbb{V}^{e_u}(u)\right] \geq \zeta\left(e_u, \kappa\right) \tag{8}$$

This formulation makes clear that a higher mismatch probability χ reduces the effective return from investing in college, thereby decreasing college attainment in equilibrium.

3.5.2 Work

During the working stage (i.e., for $j = 1, ..., J_R - 1$), an individual decides at the beginning of each period whether to be a worker or to run an entrepreneurial activity, considering her education level, her managerial ability, e_h, and the amount of financial wealth accumulated in the previous period, a. The individual then decides how much to consume and save, and, if she chooses to become an entrepreneur, whether to comply or not with business regulations and how much output to produce with capital and labor inputs, taking borrowing constraints into account.

The associated decision problem for an individual with educational status $h = \{u, s\}$ can be conveniently written in a recursive formulation with the beginning-of-period value function, $V(a, e_h)$, given by the maximum between the value of being a worker, $V^W(a, e_h)$, the value of being a formal entrepreneur, $V_f^E(a, e_h)$, and the value of being an informal entrepreneur, $V_i^E(a, e_h)$, i.e.,

$$V(a, e_h) = max\{V^W(a, e_h), V_f^E(a, e_h), V_i^E(a, e_h)\}$$

As a worker, an individual chooses how much to consume, c, and save, a', so as to maximize the continuation value under the period budget constraint, i.e.,

$$\begin{aligned} V^W(a, e_h) = & \max_{c, a' \geq 0}\{u(c) + \beta V(a', e_h)\} \\ & s.t. \quad (1+\tau_c)c + a' = (1-\tau_y)y^w + a \end{aligned}$$

where $y^w = w_h + ra$ denotes the worker's personal income, with $w_h = w_s$ for matched college-educated workers ($h = s$, $m = 0$) and $w_h = w_u$ for both non-college-educated workers ($h = u$) and mismatched college-educated workers ($h = s$, $m = 1$).

As an entrepreneur, an individual first has to decide whether to run her firm formally or informally. Compliance with business regulations allows access to credit but ensures that production activities are observable by the government, preventing registered firms from evading taxation. Consequently, the value function of being a formal entrepreneur is given by

$$\begin{aligned} V_f^E(a, e_h) = & \max_{k_f, lu_f, ls_f, c, a' \geq 0}\{u(c) + \beta V(a', e_h)\} \\ & s.t. \quad y^E = f(e_h, lu_f, ls_f, k_f) - w_s ls_f - w_u lu_f - (r+\delta)k_f + ra \\ & \qquad (1+\tau_c)c + a' = (1-\tau_y)y^E + a \\ & \qquad k_f \leq \lambda a \end{aligned}$$

where y^E denotes the formal entrepreneur's declared income, which amounts to her actual earnings.

In contrast, informal entrepreneurs evade taxation by hiding their production activities and reporting to the fiscal authorities only their capital gains, but they face a probability of detection $p(k_i)$. We assume, in addition, that (i) the government conducts audits and imposes fines after entrepreneurs have made their production decisions; and (ii) informal entrepreneurs decide how much to consume and save after observing whether they have been caught or not. Accordingly, let $V_d^E(a, e_h)$ and $V_{nd}^E(a, e_h)$ denote the informal entrepreneur's value functions corresponding to the cases of detection and non-detection, respectively. The expected value of being an informal entrepreneur can then be written as follows

$$
\begin{aligned}
V_i^E(a, e_h) = \max_{k_i, lu_i, ls_i \geq 0} & \{p(k_i)V_d^E(a, e_h) + (1 - p(k_i))V_{nd}^E(a, e_h)\} \\
s.t. \quad & k_i \leq a
\end{aligned}
$$

where the constraint on the capital rental captures that informal firms operate in a financial autarky regime. The value function in the case of non-detection is given by

$$
\begin{aligned}
V_{nd}^E(a, e_h) = \max_{c, a' \geq 0} & \{u(c) + \beta V(a', e_h)\} \\
s.t. \quad & y^E = ra \\
& (1 + \tau_c)c + a' = (1 - \tau_y)y^E + \pi + a
\end{aligned}
$$

where π represents profits from business activities, i.e.,

$$
\pi = f(e_h, lu_i, ls_i, k_i) - w_s ls_i - w_u lu_i - (r + \delta)k_i \tag{9}
$$

Note that the informal entrepreneur is able to hide profit income π from the tax authorities. However, in the case of detection, the government forces the informal entrepreneur to pay taxes on the undeclared income, augmented by the penalty surcharge factor s. Consequently, the value function of an informal entrepreneur who has been detected by the government is given by:

$$
\begin{aligned}
V_d^E(a, e_h) = \max_{c, a' \geq 0} & \{u(c) + \beta V(a', e_h)\} \\
\text{s.t.} \quad & y^E = ra \\
& (1 + \tau_c)c + a' = (1 - \tau_y)y^E + \pi + a - (1 + s)\tau_y \pi
\end{aligned}
$$

where π is defined as in Equation (9).

Finally, as in the case of workers' wages and regardless of the firm's formality status, an entrepreneur's managerial ability depends on education and the realization of mismatch risk: $e_h = e_u + \psi e_u^\epsilon$ for matched college-educated entrepreneurs ($h = s$, $m = 0$), and $e_h = e_u$ for both non-college-educated entrepreneurs ($h = u$) and mismatched college-educated entrepreneurs ($h = s$, $m = 1$).

3.5.3 Retirement

During retirement (i.e., for $j = J_R, J_R + 1, ..., J$), an individual consumes and saves on the basis of the financial wealth accumulated during the working stage. Hence, the value function of a retired individual is given as follows

$$V(a, e_h) = \max_{c, a' \geq 0} \{u(c) + \beta V(a', e_h)\}$$
$$s.t \quad (1 + \tau_c)c + a' = (1 - \tau_y)y^R + a$$

where $y^R = ra$ is the retired individual's actual income.

3.6 Equilibrium and definition of measured GDP

Let us define with $\omega = \{e_h, a, b(e_h, a), h, m\}$ the vector containing managerial ability e_h, financial wealth a, occupational status $b(e_h, a)$ (i.e., retired, worker, formal entrepreneur and informal entrepreneur), education attainment $h = h(e_u, \kappa)$, and mismatch indicator m of an agent with innate talent e_u and the idiosyncratic cost of education κ. A stationary equilibrium is given by a price vector $\{r, w_u, w_s\}$, allocations $\{c(\omega), a(\omega)\}$, occupational choices $b(e_h, a)$, decision rules for education $h(e_u, \kappa)$, unskilled and skilled labor employed by formal and informal entrepreneurs $\{lu_f(\omega), ls_f(\omega), lu_i(\omega), ls_i(\omega)\}$, formal and informal capital $\{k_f(\omega), k_i(\omega)\}$, labor and capital in the corporate sector $\{Lu_c, Ls_c, K_c\}$ and a distribution of individuals over ω, $\xi(\omega)$, such that given the risk-free interest rate r, the wage rates w_s and w_u and the tax system (i.e., s, $p(\cdot)$, τ_c and τ_y):

- The vector of policy functions $\{c(\omega), a(\omega), k_f(\omega), k_i(\omega), lu_f(\omega), ls_f(\omega), lu_i(\omega), ls_i(\omega), b(e_h, a), h(e_u, \kappa)\}$ solve the agents' decision problems described in Section 3.5.
- Capital and labor inputs $\{Lu_c, Ls_c, K_c\}$ are allocated optimally in the corporate sector.
- Capital market clears:

$$\int (k_f(\omega) + k_i(\omega)) d\xi(\omega) + K_c = \int a(\omega) d\xi(\omega)$$

- Market for skilled labor clears:

$$\int (ls_f(\omega) + ls_i(\omega))\, d\xi(\omega) + Ls_c = \int \mathbb{1}_{Ws} d\xi(\omega)$$

  where $\mathbb{1}_{Ws}$ is an indicator function that takes value of 1 if the agent is a matched college-educated worker ($h = s$, $m = 0$) and 0 otherwise.

- Market for unskilled labor clears:

$$\int (lu_f(\omega) + lu_i(\omega))\, d\xi(\omega) + Lu_c = \int \mathbb{1}_{Wu} d\xi(\omega)$$

  where $\mathbb{1}_{Wu}$ is an indicator function that takes value of 1 if the agent is either a non-college-educated worker ($h = u$) or a mismatched college-educated worker ($h = s$, $m = 1$), and 0 otherwise.

- The government budget constraint is balanced, i.e.,

$$G = \int \left\{ \tau_c c(\omega) + \tau_y y^E(\omega) + \mathbb{1}_D (1+s) \tau_y \pi(\omega) \right\} d\xi(\omega)$$

  where G are public expenditures, and $\mathbb{1}_D$ stands for an indicator function that takes a value of 1 if the individual is an informal entrepreneur who has been audited, and 0 otherwise.

- Financial intermediaries earn zero profit, i.e., Equation (3) is satisfied.

- The distribution $\xi(\omega)$ is the invariant distribution for the economy.

To close the model, we need to establish a precise definition of official−or measured −GDP. Since informal production activities are hidden from the government, official GDP does not necessarily reflect the total output of the economy. This paper adopts the assumption that official GDP is represented by total formal output, i.e.,

$$GDP = \int \mathbb{1}_f y(\omega) d\xi(\omega) + Y_c$$

where $\mathbb{1}_f$ is an indicator function taking value 1 if the agent is a formal entrepreneur and 0 otherwise, while $y(\omega)$ stands for the output of a firm operating in the entrepreneurial sector.[15] In what follows, we will likewise refer to measured TFP as total factor productivity in the formal sector.[16]

4 Calibration

We calibrate the model to the Brazilian economy. Parameter values are assigned based on either external sources or by targeting key macro and micro statistics, including estimates for the size of the informal economy, firm size distributions, college rates, and earnings premia. A description of how parameters are identified is provided below.

4.1 Identifying restrictions

Time period, life cycle duration and retirement age. The time period in the model corresponds to one year. We assume that each individual is economically active from age 20,

[15]The definition of GDP in the model is not entirely consistent with its counterpart in the data. This discrepancy arises because the Brazilian statistical institute adjusts the official GDP by incorporating estimates for the so-called non-observed economy, which includes underground production, home production, and illegal activities. Among these three components, the informal output in the model corresponds to the underground production. As discussed in Franjo *et al.* (2022), the quantitative importance of the latter may be significantly underestimated in the adjustment procedure used by the Brazilian statistical institute. For this reason, we do not include informal production in the definition of GDP in the model.

[16]Measured TFP in the model is defined as $TFP = GDP \left(\left[\mu Lu_f^{\theta} + (1-\mu) \left(\iota K_f^{\rho} + (1-\iota) Ls_f^{\rho} \right)^{\frac{\theta}{\rho}} \right]^{\frac{1-\eta}{\theta}} \right)^{-1}$, where K_f, Lu_f, and Ls_f refer to total formal capital, total formal unskilled labor, and total formal skilled labor, respectively.

retires at age 65, and lives up to 74 years. Therefore, each individual's life cycle in the model spans 55 periods ($J = 55$), with the retirement stage beginning at period $J_R = 46$. Consistent with World Bank estimates, the age of death in the model corresponds to the average life expectancy in Brazil during the 2010-2018 period, while the retirement age is set according to the Brazilian pension system.

Parameters in the inter-temporal utility function. There are two preference parameters, (β, σ). The subjective intertemporal discount factor β is chosen so that the formal capital-to-GDP ratio in the steady state equilibrium equals 2.10, which corresponds to the average capital-output ratio in Brazil during the 2004-2010 period (World Bank). The relative risk aversion coefficient σ is set to 1.5, consistent with most of the literature on occupational choice models with financial frictions (e.g. Buera *et al.*, 2011; Buera and Shin, 2013). This value also aligns with available empirical estimates for Brazil, which suggest a range for σ from 1 to 3 (Fajardo *et al.*, 2012).

Distribution of innate talents. We assume that the innate talent, e_u, is drawn from a generalized Pareto distribution with CDF:

$$\Phi(e_u) = 1 - \left(1 + \frac{\eta_p(e_u - \mu_p)}{\sigma_p}\right)^{-1/\eta_p}$$

where μ_p, σ_p and η_p are, location, scale and shape parameters, respectively. The support of the talent distribution is discretized into 20 grid points. The first point in the grid, denoted as $e_{u,min}$, is set to the value where the cumulative probability equals $\Phi(e_{u,min})$, with the latter treated as a parameter to be calibrated. The largest value corresponds to the 99.93rd percentile of $\Phi(e_u)$, while the remaining grid points are determined such that they are equidistant in probability space. We chose to calibrate $\Phi(e_u)$ to match data on the firm size distribution in the formal sector. Specifically, we use four moments to identify the four parameters $(\mu_p, \eta_p, \sigma_p, \Phi(e_{u,min}))$, targeting the share of formal firms with up to 5, 5 to 10, 11 to 20, and 21 to 50 employees. Estimates for the size distribution of formal establishments in Brazil are taken from Ulyssea (2018).

Penalty surcharge factor and probability of detection. According to Brazilian legislation, the penalty surcharge factor applied to a fraudulent taxpayer is equal to 75% of the total tax due. Therefore, we set $s = 0.75$. For the endogenous probability of detection, we follow Di Nola *et al.* (2021), Franjo *et al.* (2022) and Fernández-Bastidas (2023) by parameterizing $p(k_i)$ with a logistic function of the form

$$p(k_i) = \frac{1}{1 + p_1 exp(-p_2 k_i)}$$

where $(p_1, p_2) \in \mathbb{R}^2_+$. A property of this specification is that the probability of detection decreases with p_1. Consequently, all else being equal, the higher this parameter, the higher the expected returns from tax evasion, and thus, the larger the share of total output produced in the informal sector. We therefore set p_1 so that, in the model, the informal output to GDP ratio is equal to 37.6%. According to the Medina and Schneider (2018) estimates, this last number corresponds to the average informal output to GDP ratio in Brazil over the

1991-2015 period. To identify parameter p_2, we note that another property of the logistic specification is that the probability of detection increases quickly as capital increases, and therefore the expected gains from tax evasion decreases with the firm's size. Depending on the magnitude of parameter p_2, this property implies that informal firms optimally decide to stay small in equilibrium. A value to p_2 is therefore assigned so that share of informal firms with up to 5 workers in the model matches its empirical counterpart of 99.8% (ECINF 2003).

Production technologies and borrowing constrains. Following Allub and Erosa (2019) and Franjo *et al.* (2022), the span of control parameter η is set to 0.198. For the elasticities of substitution between capital and labor inputs, we rely on the estimates reported in Fonseca and Van Doornik (2022), setting $\theta = 0.61$ and $\rho = -0.25$.[17] This calibration implies that production technologies in our model exhibit capital-skill complementarity. The remaining share parameters in the CES production function (1), ι and μ, are set to match a wage-skill premium (w_s/w_u) of 3.55 and a share of skilled workers in the total labor force of 7.64%. As described in Section 2, these values are both computed using data from PNAD 2003. Regarding the sector-specific technology (2), we set the productivity parameter, A, such that the share of total formal capital absorbed by the corporate sector equals 0.30.[18,19] Finally, the capital depreciation rate, δ, is set to 0.06, as in Cavalcanti and Santos (2020), while the parameter controlling for the degree of financial frictions, λ, is calibrated to match a credit-to-GDP ratio of 0.42. According to the World Bank's financial structure database, this value corresponds to the average private credit-to-GDP ratio in Brazil for the period 1991-2015.[20]

Taxes. The consumption tax rate, τ_c, is set to 0.15 as in Jung and Tran (2012), who also calibrate an overlapping generations model for the Brazilian economy. The personal income tax rate, τ_y, is instead chosen to match a total fiscal revenue-to-GDP ratio of 32.4%. According to OECD revenue statistics for Latin America, this figure corresponds to the average value of Brazil's total tax revenue-to-GDP ratio over the 2000–2018 period.

Education costs, the ability-improvement function, and the probability of mismatch. To assign a value to the mismatch probability χ, we construct an occupation-based proxy for skill mismatch using PNAD 2003. We group the variable reporting occupation codes (*composição dos grupamentos ocupacionais*) into one-digit occupational categories[21] and define high-skill occupations as managers, professionals, and technicians/associate professionals. We then compute, among employed college-educated workers, the weighted share employed outside these high-skill occupations.[22] Using this definition, 75.6% of college-educated workers are

[17]Fonseca and Van Doornik (2022) report estimates for the parameters of a CES production function in each 2-digit sector of the Brazilian economy. The values we assigned to the parameters θ and ρ are unweighted averages across these sector estimates.

[18]The statistic on the capital share absorbed by Brazil's corporate sector is taken from Antunes *et al.* (2015).

[19]The fixed cost, ϕ_f, is instead determined in equilibrium by imposing a zero-profit condition in the corporate sector.

[20]The average credit-to-GDP ratio is calculated using the variable *private credit by deposit money banks and other financial institutions.*

[21]One-digit occupational categories are consistent with ISCO-88.

[22]Non-high-skill occupations include clerical support workers, services and sales workers, skilled agri-

Table 3: Calibration Results: Parameter Values

Parameters	Description	Source/ Targeted Moment	Value
(A) Externally calibrated			
σ	Relative risk aversion coefficient	Standard	1.500
δ	Capital depreciation rate	Cavalcanti and Santos (2021)	0.060
s	Surcharge factor	Brazilian law	0.750
τ_c	Consumption tax	Jung and Tran (2012)	0.150
η	Span of control	Allub and Erosa (2019)	0.198
ρ	Substitutability: capital and skilled labor	Fonseca and Van Doornik (2022)	-0.250
θ	Substitutability: capital and unskilled labor	Fonseca and Van Doornik (2022)	0.610
χ	Mismatch probability	PNAD 2003	0.244
(B) Internally calibrated			
β	Subjective discount factor	Capital-output ratio	0.962
λ	Degree of financial frictions	Credit-output ratio	1.375
$\bar{\kappa}$	Education cost: mean of idiosyncratic component	Share of educated population	7.686
σ_κ	Education cost: s.d. of idiosyncratic component	S.d. of entrepreneurial earnings	1.620
ϕ	Education cost: curvature	Entrepreneurial skill premium	1.220
ψ	Productivity jump: coefficient	Share of educated entrepreneurs by firm size	1.845
ϵ	Productivity jump: curvature	Share of educated entrepreneurs by firm size	0.958
τ_y	Income tax rate	Total fiscal revenues to GDP	0.206
A	TFP in the corporate sector	% of K used by corporations	2.220
p_1	Probability of detection	Informal output to GDP	1.4e6
p_2	Probability of detection	Size distribution informal firms	5.546
μ_p	Location Pareto distribution	Size distribution formal firms	2.743
σ_p	Scale Pareto distribution	Size distribution formal firms	1.455
η_p	Shape Pareto distribution	Size distribution formal firms	0.228
$\Phi(e_{min})$	Probability mass in the minimum ability	Size distribution formal firms	0.471
μ	Weight of unskilled labor in production	Wage skill premium	0.464
ι	Weight of capital in production	Share of educated workers	0.709

employed in high-skill occupations, while the remaining 24.4% are employed in non-high-skill occupations. We set $\chi = 0.244$ accordingly.

To parameterize the cost of attending college, we follow Heathcote *et al.* (2010) by assuming a lognormal distribution for the idiosyncratic component κ, i.e.,

$$ln(\kappa) \sim \mathcal{N}(\bar{\kappa},\, \sigma_\kappa^2)$$

We set the mean $\bar{\kappa}$ to match a population college-education rate of 8.36%, based on PNAD 2003. As for the standard deviation σ_κ, we note that in the model, this parameter controls the dispersion of educational costs among individuals with the same innate talent. Consequently, it also influences the overall dispersion of population earnings conditional on education. We take advantage of this property and identify the parameter σ_κ by matching a measure of earnings volatility. Specifically, we use the standard deviation of earnings among Brazilian entrepreneurs, which according to the PNAD 2003 survey is equal to 1.053. Information on population earnings is also useful to assign a value to the curvature parameter ϕ. According

cultural, forestry and fishery workers, craft and related trades workers, plant and machine operators and assemblers, and elementary occupations.

Table 4: Calibration Results: Targeted Moments

Moments	Source	Data	Model
Capital-Output ratio (formal)	Allub and Erosa (2019)	2.100	2.109
Credit-Output ratio	World Bank Database	0.420	0.410
Informal output to GDP	Medina and Schneider (2018)	0.376	0.345
% of K used by corporations	Antunes et al. (2015)	0.300	0.288
Total fiscal revenues to GDP	OECD revenues statistics	0.320	0.325
Share of educated individuals	PNAD 2003	0.084	0.090
Share of educated workers	PNAD 2003	0.076	0.074
S.d. entrepreneurial earnings	PNAD 2003	1.053	1.099
Entrepreneurial skill premium	PNAD 2003	3.927	3.920
Wage skill premium	PNAD 2003	3.546	3.540
Share of educated entrepreneurs by firm size			
6-10 workers	PNAD 2003	0.285	0.214
$>$ 10 workers	PNAD 2003	0.368	0.423
Size distribution: informal firms			
$\leq$ 5 workers	ECINF 2003	0.998	0.975
Size distribution: formal firms			
$\leq$ 5 workers	Ulyssea (2018)	0.701	0.676
6-10 workers	Ulyssea (2018)	0.141	0.106
11-20 workers	Ulyssea (2018)	0.083	0.156
21-50 workers	Ulyssea (2018)	0.048	0.063

to Equation (7), for a given κ, the cost of attending college decreases with ϕ for any innate talent e_u. Hence, all else being equal, the higher this parameter, the larger the share of college-educated individuals, and, in equilibrium, the lower the earnings skill premia. Consequently, we calibrate ϕ by matching the college earnings premium for entrepreneurs. In the data, we compute this as the ratio of the average earnings of college-educated entrepreneurs to the average earnings of non-college-educated entrepreneurs. As described in Section 2, we estimate the entrepreneurial skill premium to be 3.93. Finally, the specification of the improvement-ability function (6) introduces two additional education-related parameters, (ϵ, ψ). We calibrate these by targeting the share of educated entrepreneurs in firms with 6 to 10, and more than 10 employees (PNAD 2003). Since college education boosts managerial productivity, in the model, educated entrepreneurs tend to be concentrated in medium- and large-sized enterprises. Consequently, the targeted moments are reasonably informative in identifying parameters ϵ and ψ.

In summary, our calibration strategy partitions the model's parameters into two subvectors: one containing those that are fixed according to external sources (i.e., σ, s, η, θ, ρ, δ, τ_c, χ), and another containing the internally calibrated parameters (i.e., β, μ_p, σ_p, η_p, $\Phi(e_{u,\min})$, p_1, p_2, ι, μ, A, λ, τ_y, $\bar{\kappa}$, σ_κ, ϕ, ϵ, ψ). Values for the latter are jointly assigned by minimizing a loss function that computes the distance between the model's predicted moments and the targeted moments from the data. The results are provided in Tables 3, 4, and 5. Table 3 reports the values of the calibrated parameters, while Tables 4 and 5 provide a comparison between the model and the data for targeted and non-targeted moments, respectively.[23]

[23]We have also tested more formally whether the calibration restrictions described above allow for the

Table 5: Calibration Results: Non-targeted Moments

Moments	Source	Data	Model
(A) *Share of non-college-educated entrepreneurs*:			
All firms	PNAD 2003	0.898	0.915
In formal firms	ECINF 2003	0.826	0.800
In informal firms	ECINF 2003	0.937	0.961
(B) *Share of informal workers*:			
All workers	PNAD 2003	0.269	0.282
Unskilled workers	PNAD 2003	0.277	0.289
Skilled workers	PNAD 2003	0.171	0.195
(C) *Formal earnings premium*:			
Entrepreneurs	ECINF 2003	2.440	2.098
Workers	ECINF 2003	1.386	1.091
Share of informal firms	Ulyssea (2018)	0.698	0.576
Gini index for wealth	Franjo *et al.* (2022)	0.784	0.788
Average capital-output ratio across informal firms	Erosa *et al.* (2023)	1.040	0.931

4.2 Goodness of fit and external validation of the model

As Table 4 illustrates, the model replicates the targeted moments quite well. In particular, it accurately captures the earnings skill premia for both workers and entrepreneurs, and reasonably matches the share of college-educated entrepreneurs by firm size. The model also does a good job of replicating the size difference between formal and informal firms, with implied aggregate informal output accounting for around 35% of Brazilian measured GDP, in line with the data. Similarly, all other targeted aggregate ratios in the model's steady-state equilibrium are close to their empirical counterparts.

Regarding non-targeted moments, Table 5 shows that the model performs fairly well in capturing all the considered dimensions of the data. In particular, the model is consistent with the data in predicting that (i) informality is more prevalent among less educated entrepreneurs (see Panel A); (ii) the informal sector is more intensive in unskilled labor; and (iii) informal production is less capital intensive than formal production (see capital-output ratios in Tables 4 and 5). These findings highlight that combining an endogenous probability of detection with capital-skill complementarity in a heterogeneous-agent model of entrepreneurship with financial frictions allows the model to reproduce the differences between formal and informal firms discussed in Section 2, both in terms of workers' skill intensity and managers' educational attainment.[24] In this respect, it is also worth highlighting that the model accurately captures the main features of the formal earnings premia in the data. Although it slightly underestimates their magnitude, the model predicts—consistent with

identification of parameters. Details on the procedure and results are provided in the Online Technical Appendix B.

[24]Intuitively, the calibration of the endogenous probability of detection implies that informal firms are predominantly smaller than formal firms. As shown in Tables 4 and 5, this is reflected in a lower average capital-output ratio among informal firms (i.e., 0.93 vs. 2.11). Due to capital-skill complementarity, lower capital intensity generates a relatively higher demand for unskilled labor in the informal sector, even though firms have access to the same production technology. Regarding the educational attainment of entrepreneurs, in the next section, we will show that the difference between formal and informal firms in the model arises from the self-selection of talent into entrepreneurship, driven by the wage-skill premium and the boost in managerial skills induced by college education.

the data—that operating in the formal sector yields earnings gains for both entrepreneurs and workers, with the premium for the latter being approximately 1.9 times higher than that for the former (see Panel C).

5 Results

In this section, we first characterize the properties of the calibrated model (i.e., the benchmark economy), with a special focus on occupational choices, human capital accumulation, and informality. We then investigate how individual access to college education and firms' access to credit interact with the informal sector and shape aggregate economic performance, identifying the channels through which these factors affect individual and sectoral outcomes.

Throughout the analysis, we distinguish educational attainment from realized occupational skills. Accordingly, non-college individuals ($h = u$) and mismatched college-educated individuals ($h = s, m = 1$) are classified as unskilled, whereas matched college-educated individuals ($h = s, m = 0$) are classified as skilled. Thus, mismatched college-educated individuals are treated as unskilled in labor-market and entrepreneurial decisions, while still being counted as college-educated in education statistics.

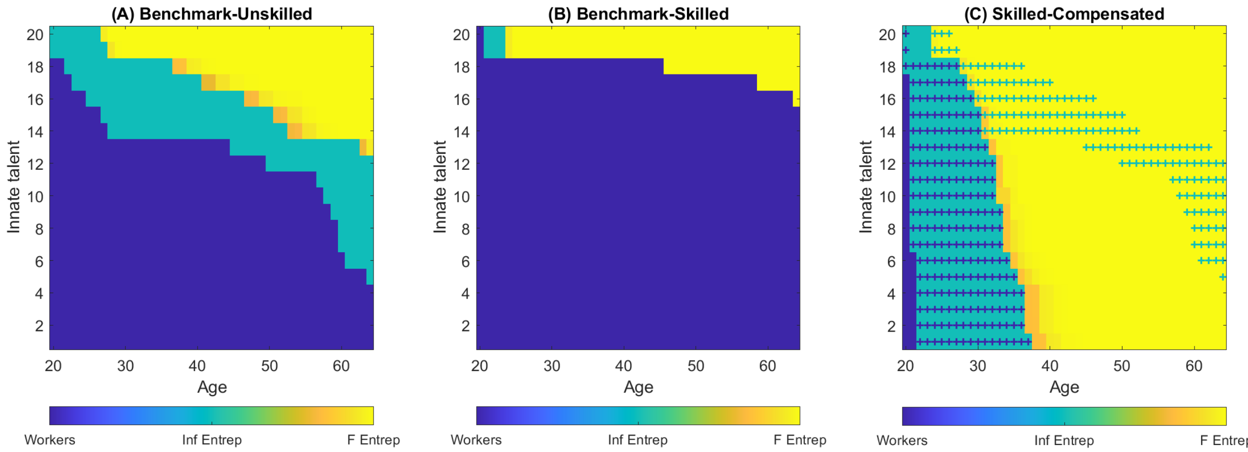
5.1 Drivers of firm informality in the benchmark economy

For each level of innate talent, Figure 3 illustrates the occupational maps across the working life for individuals in the benchmark economy, differentiating between unskilled (Panel A) and skilled (Panel B) agents. The reported occupational patterns reveal substantial behavioral heterogeneity conditional on realized occupational skills. In particular, the figure shows that informality is predominantly concentrated among firms run by unskilled entrepreneurs. While this observation is in line with the findings highlighted in the previous section regarding the model's ability to replicate key features of the data, the analysis of occupational maps provides additional insights into the underlying mechanisms. Specifically, Figure 3 shows that most of the unskilled individuals operate as informal entrepreneurs for a portion of their working lives. In contrast, only the most talented skilled managers find it optimal to run their enterprises informally for a brief period before complying with business regulations. All other matched college-educated individuals either choose to work as skilled employees before transitioning directly to formal entrepreneurship or remain skilled workers throughout their entire working lives.

In the benchmark economy, this heterogeneity in occupational choices is driven by the combined effects of two channels: the wage-skill premium and the entrepreneurial human capital accumulation associated with college education, which influence the self-selection of talent into entrepreneurship. To see this, first recall that, in the calibrated model, skilled individuals benefit from higher wages, which (i) make paid employment the optimal occupational choice for agents in the low-to-middle range of the innate talent distribution and (ii) facilitate more efficient asset accumulation for individuals with high managerial ability before they transition directly to formal entrepreneurship.[25] This latter path is not optimal for unskilled formal entrepreneurs because the unskilled wage is too low, leading them

[25] In our model, age plays a crucial role in determining occupational choices, as it captures the gradual accumulation of assets over time. Panels A and B of Figure 3 show that nearly all individuals begin their careers as workers and, as they age, some transition into entrepreneurship depending on managerial ability.

Figure 3: Occupational maps



Notes: The figure illustrates occupational choices over the life cycle for: (A) unskilled individuals, (B) skilled individuals, and (C) matched college-educated individuals whose labor income is adjusted to offset the wage-skill premium. Colors indicate occupational status: blue for workers, turquoise for informal entrepreneurs, and yellow for formal entrepreneurs. In Panel C, blue markers identify ages at which individuals with the same innate ability operate as skilled workers in the benchmark equilibrium and as informal entrepreneurs in the compensated experiment. Similarly, turquoise markers identify ages at which individuals with the same innate ability operate as formal entrepreneurs in the compensated experiment and as unskilled informal entrepreneurs in the benchmark equilibrium.

to prefer starting their firms informally and engaging in tax evasion as a more effective means of accumulating financial wealth to overcome borrowing constraints.[26] Depending on the innate talent, Panel A of Figure 3 shows that, for these individuals, engaging in informal entrepreneurship can either occur immediately—as in the case of the highly talented entrepreneurs—or later in life, after operating for a certain period as unskilled workers.[27]

The low unskilled wage also accounts for the occupational switches observed among low-ability unskilled individuals. Unlike their skilled counterparts with the same innate abilities, who remain skilled workers throughout their careers, these individuals face a lower opportunity cost of entrepreneurship. Hence, once they have accumulated enough wealth to self-finance a small firm, operating informally becomes more profitable than remaining in

This pattern arises because all agents start their working lives with zero assets and face financial frictions; together, these features make paid employment the optimal initial occupation for many individuals, allowing them to accumulate sufficient wealth to finance entrepreneurial activity later in life.

[26]Panel A of Figure 3 shows that early entry into informal entrepreneurship is concentrated among unskilled individuals with the highest managerial ability, who have strong incentives to operate a firm despite limited initial wealth. Because financial frictions constrain the scale at which they can operate formally, while the probability of detection is low for small firms with little capital, the gains from tax evasion may initially outweigh the benefits of formalization and access to credit, making delayed formalization the optimal path for these entrepreneurs. The same logic also explains why the most talented skilled individuals, despite facing a high skilled wage, start their businesses informally rather than remaining workers for a longer period (see Panel B of Figure 3).

[27]The outcome that the highest-ability unskilled individuals optimally start their working lives directly as informal entrepreneurs is driven by the assumption of a CES production technology (see Equation 1), which permits positive production even with zero capital.

Table 6: Firm dynamics in the calibrated model

(A) *Relative size among entrants*	Ratio
(1) **Formal firms run by skilled entrepreneurs**:	
Number of workers	2.92
Capital	1.60
Sales	2.64
(2) **Informal firms run by skilled entrepreneurs**:	
Number of workers	6.11
Capital	0.71
Sales	5.06
(B) *Average firm growth*	Percentage
(1) **Formal firms**:	
Skilled entrepreneurs	2.2
Unskilled entrepreneurs	1.5
(2) **Informal firms**:	
Skilled entrepreneurs	6.7
Unskilled entrepreneurs	2.8

Note: In Panel A, each size measure is reported as a ratio relative to the average value for entrant firms managed by unskilled entrepreneurs. In Panel B, average firm growth rates are calculated as the mean of annual percentage changes in firm output.

paid employment, even when managerial ability is limited. This mechanism explains why some unskilled agents enter informal entrepreneurship without ever formalizing.

While the previous discussion emphasizes the role of labor-income differentials, college education also affects firm informality through entrepreneurial human capital accumulation. In the calibrated model, schooling enhances managerial productivity of matched individuals (i.e., $\psi > 0$), leading firms run by skilled entrepreneurs to enter at a larger scale and grow faster than those managed by unskilled business owners. This property is evident in Table 6, which reports various statistics characterizing firm dynamics in the benchmark economy. Panel A shows that, for all reported measures of size, entrant firms run by skilled entrepreneurs are, on average, larger than those operated by unskilled individuals, regardless of whether the businesses are formalized. Moreover, Panel B shows that firms run by skilled entrepreneurs grow faster over the life cycle.[28] Since the probability of detection, $p(k_i)$, increases with a firm's capital size, these differences in firm dynamics imply that the expected returns from informality are generally lower for matched college-educated entrepreneurs, giving them a stronger incentive to formalize their businesses. This is the second channel through which human capital accumulation affects firm informality in the model.

One way to disentangle these two channels is to examine how the occupational choices of matched college-educated individuals would change if their labor income advantage were removed. Specifically, we keep factor prices, r_k, w_u, and w_s, as well as college rates across innate abilities, fixed at their benchmark equilibrium values, and recompute the occupational choices of skilled individuals after artificially reducing their disposable labor income to offset the wage-skill premium. This exercise allows us to isolate the role of labor-income differentials by comparing the resulting occupational maps with those of skilled individuals in the benchmark equilibrium. At the same time, by comparing the *compensated* occupational maps with those of unskilled individuals in the benchmark economy, we can assess the partial contribution of entrepreneurial human capital accumulation to firm formalization.

[28] Differences in firm dynamics conditional on the entrepreneur's level of education predicted by the model are consistent with the evidence documented in Queiró (2021).

The results of the experiment are reported in Panel C of Figure 3. Offsetting the wage-skill premium causes all matched college-educated individuals to engage in entrepreneurship over the life cycle. In particular, these agents follow the same broad pattern as unskilled individuals who eventually become formal entrepreneurs: they start their businesses informally before transitioning to formal entrepreneurship. This finding provides two main insights. First, comparing Panel C with Panel B shows that college education reduces firm informality indirectly through the wage-skill premium. In the benchmark equilibrium, the higher skilled wage raises the opportunity cost of entrepreneurship, inducing most matched college-educated individuals to remain skilled workers rather than enter informal entrepreneurship. When this labor-income advantage is offset, these individuals instead choose informal entrepreneurship over part of their life cycle. This indirect effect is illustrated in Panel C by the area marked with blue crosses, which identifies points where individuals are skilled workers in the benchmark equilibrium but informal entrepreneurs in the compensated experiment.

Second, comparing Panel C with Panel A shows that, for a given level of innate ability, once labor-income differentials are offset, skilled individuals who engage in entrepreneurship formalize their businesses much earlier over the firm life cycle than unskilled entrepreneurs in the benchmark equilibrium. This difference reflects the role of entrepreneurial human capital accumulation in shaping firm dynamics: firms run by skilled entrepreneurs enter at a larger scale and grow faster than those managed by unskilled entrepreneurs. As a result, the probability of detection increases more quickly for matched college-educated entrepreneurs, allowing them to reach earlier in the firm life cycle the point at which the benefits of formalization associated with access to credit outweigh the expected gains from tax evasion. Put differently, the lack of entrepreneurial human capital accumulation increases the persistence of firm informality over the life cycle. Since the benchmark economy features a large share of unskilled individuals, this mechanism contributes substantially to the persistence of firm informality in equilibrium, as shown by the area marked with turquoise crosses in Panel C.

To summarize, college education, when successfully translated into occupational skills, reduces firm informality through two channels: indirectly, by increasing the opportunity cost of entrepreneurship through the wage-skill premium; and directly, by increasing entrepreneurial human capital, which raises firm scale and growth, increases the probability of detection, and reduces the expected returns from tax evasion. In the next section, we assess how these mechanisms shape aggregate outcomes through a series of counterfactual experiments.

5.2 The role of human capital in the aggregate economy

To study the aggregate role of human capital, we bring the college rate in the calibrated benchmark economy up to the level observed in Brazil in 2012, mimicking an educational expansion the country has already undergone. Between 2003 and 2012, Brazil's college attainment rate rose from 8.4% to 12.5% (PNAD 2003; PNAD-C 2012), a period during which the country also experienced a marked decline in labor informality and a recovery in aggregate productivity. Using a search-and-matching model of the Brazilian labor market, Haanwinckel and Soares (2021) show that this shift in workforce composition was a primary driver of the decline in salaried informality over the same period, while Cox (2024) documents that the resulting expansion in the supply of college-educated workers led firms to grow larger.

Motivated by this evidence, we lower the average cost of tertiary education in the model (i.e., parameter $\bar{\kappa}$) until the college rate reaches its 2012 level, while keeping all other parameters fixed at their calibrated values.[29] Comparing the resulting steady-state equilibrium with the benchmark allows us to quantify the aggregate consequences of an educational expansion that Brazil has already experienced.

The results are presented in Table 7 and Figure 4. Table 7 reports several key statistics for the benchmark economy (first column) and the higher-human-capital scenario with a college rate of 12.5% (second column). Figure 4 illustrates the distribution of individuals across occupations by innate ability, distinguishing between skilled (first column) and unskilled individuals (second column) in the benchmark economy (panels A and B) and in the counterfactual with cheaper access to college education (panels C and D).

5.2.1 The informal economy

As shown in Table 7, increasing the proportion of college-educated individuals in the long-run equilibrium leads to a substantial reduction in the size of the informal economy, with the ratio of informal output to official GDP falling from 34.5% to 23.2%. This outcome is driven by the combined effects of two forces that affect both the numerator and the denominator of this ratio. On the one hand, in the counterfactual scenario with higher human capital, both the proportion of informal entrepreneurs in the population and the share of informal firms decline significantly, leading to a 21% reduction in aggregate informal production. On the other hand, despite the share of formal entrepreneurs in the population increasing by only 1.2 percentage points, the model predicts that raising the population's college rate to the level observed in Brazil in 2012 increases official GDP by 7.7%. Together, these two margins reduce the size of the informal economy by approximately 11.3 percentage points.

To shed light on the mechanisms underlying these findings, Table 7 presents several statistics that offer insights into the general equilibrium effects of human capital accumulation. Specifically, the table shows that, in the counterfactual economy where the average cost of college education is lower, the share of skilled workers in the population increases, while the share of unskilled workers decreases relative to the benchmark economy. This shift in the human capital composition of the labor force results in a lower skilled wage and, simultaneously, a higher unskilled wage in equilibrium, leading to a 13.7% reduction in the wage-skill premium.[30] Consequently, the opportunity cost of entrepreneurship rises for unskilled individuals while falling for matched college-educated individuals. Put differently, in general equilibrium, the adjustment in the wage structure induced by a higher college rate encourages entrepreneurship among matched college-educated individuals while incentivizing paid employment among unskilled ones. This pattern is clearly illustrated in Panels A-D

[29] We follow a supply-driven approach, consistent with this evidence: the expansion of college education in Brazil over this period has been largely attributed to the 1996 higher education reform, which lowered entry barriers for private colleges (Cox, 2024), and similar supply-driven dynamics have been documented in other Latin American countries (Firpo and Portella, 2019). This contrasts with demand-driven explanations—such as skill-biased technological change—which are more relevant for developed economies, where they have been linked to rising educational attainment and increasing wage dispersion (e.g., Autor, 2014; Van Reenen, 2011).

[30] There is substantial evidence documenting that changes in relative labor supply, driven by higher college rates, played a critical role in the decline of the wage-skill premium observed in several developing countries in Latin America during the 2000s (see, e.g., Soares and Haanwinckel, 2017; Fernández and Messina, 2018; Haanwinckel and Soares, 2021). The predictions of our model are consistent with this evidence.

Table 7: The role of human capital

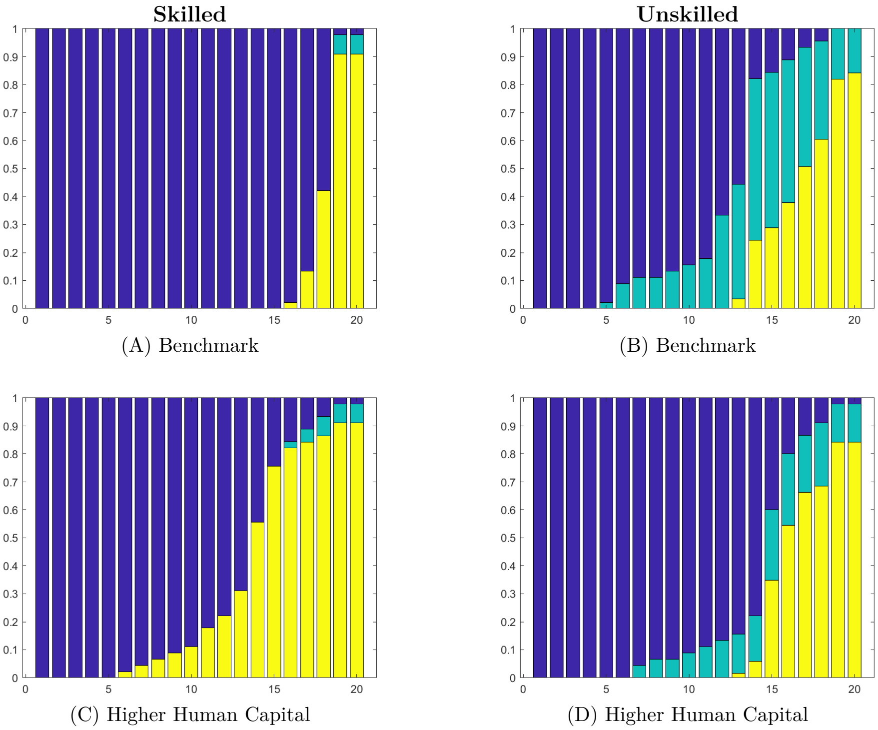
Statistic	Benchmark	Higher Human Capital	No Entrep Human Capital
Education metrics			
College rate (% of population)	9.0%	12.5%	11.1%
Informality metrics			
Size informal economy (% of official GDP)	34.5%	23.2%	32.2%
Informal firms (% of total firms)	57.6%	48.7%	50.9%
Workers informality (% of total workers)	28.2%	21.1%	27.1%
Changes relative to benchmark (%)			
(i) Macro aggregates and prices			
Informal production		-21.0%	-17.8%
Official GDP		+7.7%	+3.9%
Measured TFP		+2.3%	+1.1%
Skilled wage		-12.2%	-16.8%
Unskilled wage		+3.0%	+3.5%
Wage skill premium		-13.7%	-18.2%
Entrepreneurial skill premium		-3.0%	+4.8%
(ii) Fiscal metrics			
Fiscal revenues		+8.6%	+6.2%
Tax evasion		-25.8%	-20.1%
Entrepreneurship rates (%)			
Total entrepreneurs in population	19.1%	18.2%	21.7%
Entrepreneurs among skilled individuals	12.5%	15.6%	3.2%
Entrepreneurs among unskilled individuals	19.6%	18.4%	23.4%
Formal entrepreneurs (% of population)	8.1%	9.3%	10.7%
Informal entrepreneurs (% of population)	11.0%	8.9%	11.1%
College-educated formal entrepreneurs (% of formal entrepreneurs)	14.7%	21.2%	6.7%
Workers shares (% of population)			
Total workers	80.9%	81.8%	78.3%
Skilled workers	5.9%	8.0%	8.1%
Unskilled workers	74.9%	73.8%	70.1%

of Figure 4, which show that, in the equilibrium with higher human capital accumulation, entrepreneurship becomes more prevalent among matched college-educated individuals, while the distribution of unskilled agents becomes more concentrated in paid employment. These effects are also evident in Table 7, which reveals that the observed decline in the overall entrepreneurship rate (from 19.1% to 18.2%) masks a heterogeneous response by realized occupational skills: the rate drops among unskilled individuals, from 19.6% to 18.4%, while it increases significantly among skilled individuals, rising from 12.5% to 15.6%.

These opposing effects on entrepreneurship rates contribute to reducing the size of the informal economy through two main channels. On the one hand, since informality in the benchmark economy is predominantly concentrated among unskilled entrepreneurs, the decline in entrepreneurship within this group directly reduces firm informality. Panels B and D of Figure 4 illustrate this effect: in the counterfactual economy where the unskilled wage is higher, informal entrepreneurship declines across all levels of innate talent. By reducing the share of firms operating informally, from 57.6% to 48.7%, this effect contributes to the decline in informal production observed in our experiment.

On the other hand, the entrepreneurial human capital channel discussed in Section 5.1 implies that the increase in entrepreneurship among skilled individuals, driven by the general-equilibrium decline in skilled wages, promotes the creation of optimally larger firms with

Figure 4: Distribution of skilled and unskilled individuals across occupations.



(A) Benchmark

(B) Benchmark

(C) Higher Human Capital

(D) Higher Human Capital

Notes: The figure shows the distribution of individuals across occupations by innate ability, distinguishing between skilled (left column) and unskilled individuals (right column). Yellow, turquoise, and dark blue represent the share of formal entrepreneurs, informal entrepreneurs, and workers, respectively. Panels A and B depict the benchmark economy. Panels C and D represent the counterfactual scenario with cheaper access to college education ("Higher Human Capital").

stronger incentives to operate formally.[31] This effect is illustrated in Panels A and C of Figure 4: in the counterfactual economy, formal entrepreneurship among skilled individuals increases across all levels of innate ability relative to the benchmark equilibrium. By boosting formal production, these distributional shifts further contribute to the decline in the size of the informal sector.

Beyond these occupational responses, higher college attainment also reduces the extensive margin of informality through two additional forces that strengthen formalization incentives. First, in our model, formal firms are larger than informal firms and, as a result, rely more heavily on skilled labor due to capital-skill complementarity. Consequently, the decline in the wage-skill premium observed in the counterfactual equilibrium lowers the relative cost of formal production, strengthening entrepreneurs' incentives to formalize their businesses, particularly among agents with high managerial ability. This general-equilibrium effect explains why, in our experiment, despite the increase in the unskilled wage, formalization increases not only among matched college-educated individuals but also among unskilled entrepreneurs with high innate talent, as illustrated in Panels B and D of Figure 4.

Second, by increasing the number of individuals for whom college education translates

[31] Empirical support for this mechanism is provided by Cox (2024), who, using Brazilian data, documents that an increase in college attainment within the population leads to a rise in the total number of large firms.

into occupational skills, the higher college rate expands the mass of potential entrepreneurs with enhanced managerial abilities. As discussed in Section 5.1, these entrepreneurs accumulate more entrepreneurial human capital than unskilled business owners and therefore face lower expected returns from tax evasion when operating a firm. By fostering formalization among this group, this composition effect further contributes to reducing the size of the informal economy.

Beyond their impact on aggregate informal production and official GDP, the importance of the above channels in promoting formal entrepreneurship is clearly illustrated in Table 7. The share of formal entrepreneurs in the population increases from 8.1% to 9.3%, while the share of informal entrepreneurs declines from 11% to 8.9%. Moreover, the proportion of college-educated individuals among formal entrepreneurs rises substantially, from 14.7% in the benchmark economy to 21.2% in the counterfactual scenario, highlighting the critical role of entrepreneurial human capital in reducing the extensive margin of informality. These changes imply a lower share of individuals under-reporting income to fiscal authorities, leading to a 25.8% reduction in tax evasion. Coupled with the expansion of formal production, this decline generates substantial fiscal gains, with total government revenues increasing by 8.6% in the long run.[32]

5.2.2 Official GDP and measured TFP

An interesting result of the above experiment is the significant impact of higher human capital on measured GDP in Brazil, which increases by 7.7%. This increase reflects the formal-sector reallocation generated by the mechanisms discussed above: matched college-educated individuals become more likely to enter entrepreneurship, while unskilled entrepreneurs, especially those with high ability, face stronger incentives to formalize their activities. Together, these mechanisms shift production from the informal to the formal sector. The labor market adjusts accordingly, with employment reallocating toward formal firms and the share of informal workers declining from 28.2% in the benchmark economy to 21.1% in the counterfactual scenario. This prediction is consistent with the evolution of labor informality in Brazil, where informality rates among salaried workers declined from 28.1% in 2003 to 17.3% in 2012 (Haanwinckel and Soares, 2021). The model can account for roughly two thirds of this decline, providing additional external validation for the mechanisms linking human capital accumulation to the formalization of labor and production.

The experiment also has non-trivial implications for measured TFP, as higher human capital affects aggregate productivity through several margins that operate in different directions. On the one hand, there is a substantial increase in formal entrepreneurship among skilled individuals, which boosts aggregate productivity. On the other hand, this expansion is accompanied by a decline in the average ability of skilled entrepreneurs, as shown by comparing the yellow areas in Panels A and C of Figure 4. In contrast, Panels B and D show that formal entrepreneurship among unskilled individuals becomes more concentrated at higher levels of innate ability, as higher unskilled wages raise the opportunity cost of en-

[32]From a fiscal-policy perspective, these results suggest that reforms facilitating access to college education may contribute to reducing tax evasion and firm informality in developing economies. In this respect, our analysis complements and extends the findings of Soares and Haanwinckel (2017), who argue that improvements in the schooling and skill composition of the labor force can be an effective policy lever for reducing labor informality.

trepreneurship for lower-ability individuals, prompting them to remain in paid employment rather than enter entrepreneurship. As reported in Table 7, the net effect of these opposing forces is a 2.3% increase in measured TFP. This result is broadly consistent with evidence that Brazil experienced a recovery in TFP during the 2000s. Ferreira and Veloso (2013), for instance, document a reversal in Brazil's TFP dynamics after 2003, following a decline during the previous decade.

5.2.3 Entrepreneurial human capital

In this section, we isolate the quantitative importance of entrepreneurial human capital accumulation in driving the results discussed above by conducting a counterfactual experiment. Specifically, we reassess the effects of improving access to college—by reducing the parameter $\bar{\kappa}$—under the assumption that tertiary education does not enhance managerial ability, implemented by setting $\psi = 0$. The contribution of entrepreneurial human capital can then be isolated by comparing the resulting steady-state equilibrium to that of the higher-human-capital economy, which features both improved access to college and enhanced managerial ability for matched individuals (i.e., $\psi > 0$).[33]

The results, reported in the third column of Table 7, show that entrepreneurial human capital is central to the aggregate gains from higher educational attainment. Absent this channel, the economy captures only about half of the official GDP gain (3.9% vs. 7.7%) and of the increase in measured TFP (1.1% vs. 2.3%) achieved in the higher human capital economy. The decline in firm informality is also substantially milder: the informal sector shrinks by only 2.3 percentage points, from 34.5% to 32.2%, compared to an 11.3 percentage point decline, to 23.2%, when entrepreneurial human capital is present.[34]

The intuition follows the same logic as in Section 5.2: a lower average cost of college still shifts the skill composition of labor supply, reducing the wage-skill premium by 18.2% and, through capital-skill complementarity, raising average firm size and formalization incentives—reinforced by the rise in the unskilled wage (+3.5%), which raises the relative cost of informal production. However, without a return to managerial ability, college now yields returns only through the wage-skill premium. As a result, the entrepreneurship rate among skilled individuals rises only to 3.2% (compared to 15.6% with entrepreneurial human capital), while it remains elevated among unskilled individuals (23.4% vs. 18.4%), concentrating entrepreneurship exactly where firm informality is most prevalent.

This composition effect leads to a substantially milder decline in informal activity, which falls by only 17.8%, compared to 21.0% in the higher-human-capital scenario. Similarly, the share of informal firms declines to 50.9%, as opposed to 48.7% when entrepreneurial human capital is present, and the share of informally employed workers falls only to 27.1% of total workers compared to 21.1%, in the higher-human-capital economy. As a result, the proportion of individuals under-reporting income to fiscal authorities also decreases less sharply, resulting in a smaller reduction in tax evasion (20.1% vs. 25.8%). These effects

[33] To facilitate comparison between the two frameworks, in the model without entrepreneurial human capital accumulation we set $\bar{\kappa}$ to the same level used in the counterfactual scenario with $\psi > 0$. All remaining parameters are kept fixed at their calibrated values.

[34] All changes in the third column of Table 7 are computed relative to a counterfactual economy with $\psi = 0$ and $\bar{\kappa}$ set to its benchmark value, in order to isolate the effect of improved access to college education in an economy without entrepreneurial human capital.

are further reinforced by the much lower share of educated formal entrepreneurs among all formal entrepreneurs—just 6.7% in the absence of entrepreneurial human capital, compared to 21.2% with it.

Altogether, these results show that entrepreneurial human capital is a central driver of the aggregate gains from education, for both firm informality and official measures of output and productivity. This finding points to cross-country differences in entrepreneurial human capital accumulation as a potentially important factor behind disparities in income per capita and aggregate productivity.

5.3 Interaction between human capital and credit

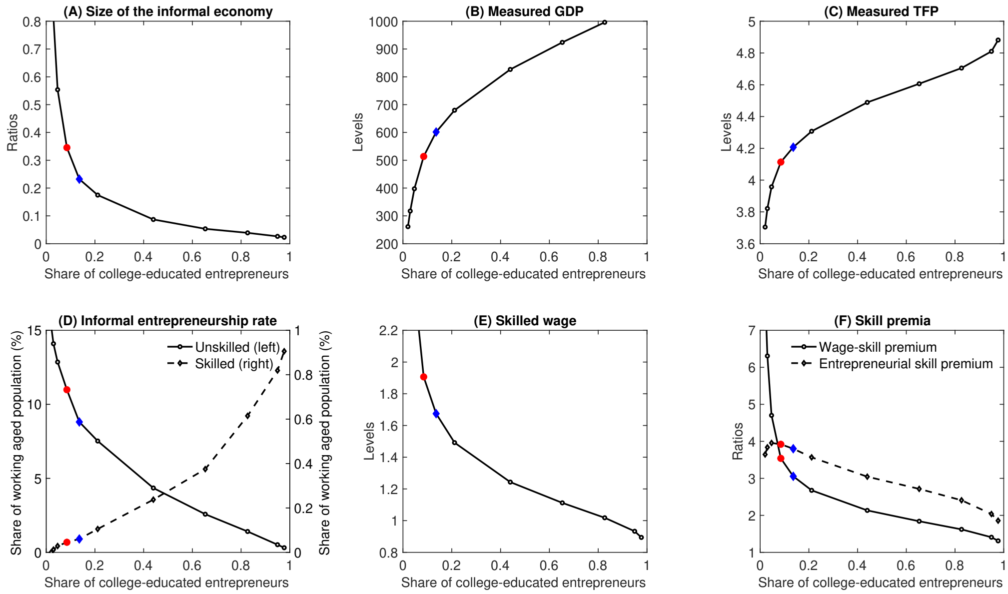
The interaction between human capital accumulation and access to credit is particularly important for understanding the occupational choices of matched college-educated individuals with high entrepreneurial ability. In this respect, Panel C of Figure 4 shows that the increase in human capital accumulation induces more of these agents to enter business ownership, but part of this response takes the form of informal entrepreneurship. As a result, despite the contraction in aggregate informal production, we find that the mass of skilled informal entrepreneurs increases in the economy with better access to college education. This pattern is concentrated among high-ability skilled entrepreneurs because, as discussed in Section 5.1, these agents have stronger incentives to start businesses, but financial frictions may still prevent them from reaching, early in the firm life cycle, the scale at which formalization becomes optimal. This constraint plays a more prominent role in the counterfactual economy because the general-equilibrium decline in the skilled wage increases the incentives of matched college-educated individuals to enter entrepreneurship.

To evaluate the robustness of these predictions, we conduct an additional experiment by solving for the long-run equilibrium under alternative values of the parameter governing the average cost of education, $\bar{\kappa}$, while holding all other parameters at their calibrated values. Figure 5 presents the results, illustrating how key equilibrium outcomes respond to changes in the average cost of education. To facilitate comparison with the data discussed in Section 2, the figure plots steady-state values against the share of college-educated entrepreneurs, which increases monotonically as $\bar{\kappa}$ declines.

Two main results emerge. First, Figure 5 generalizes the insights from our counterfactual exercise, highlighting a robust negative relationship in the model between the size of the informal economy and the share of college-educated entrepreneurs. This result is consistent with the cross-country evidence documented in Section 2, which shows that countries characterized by lower educational attainment among entrepreneurs also tend to have larger informal sectors (see Figure 2). Second, and most importantly, the experiment shows that increasing college attainment alone is not sufficient to eliminate firm informality (see panel A), which persists even in an economy where nearly all entrepreneurs are college-educated. This persistence is driven by the rising rate of informal entrepreneurship among skilled individuals as the cost of education declines (panel D). Together with the monotonic decline in skilled wages (panel E), this finding complements and extends the discussion above: when educational attainment becomes widespread, financial frictions remain a central force sustaining firm informality.

To further isolate the role of financial frictions, we repeat this comparative-statics ex-

Figure 5: Cost of college education and the aggregate economy

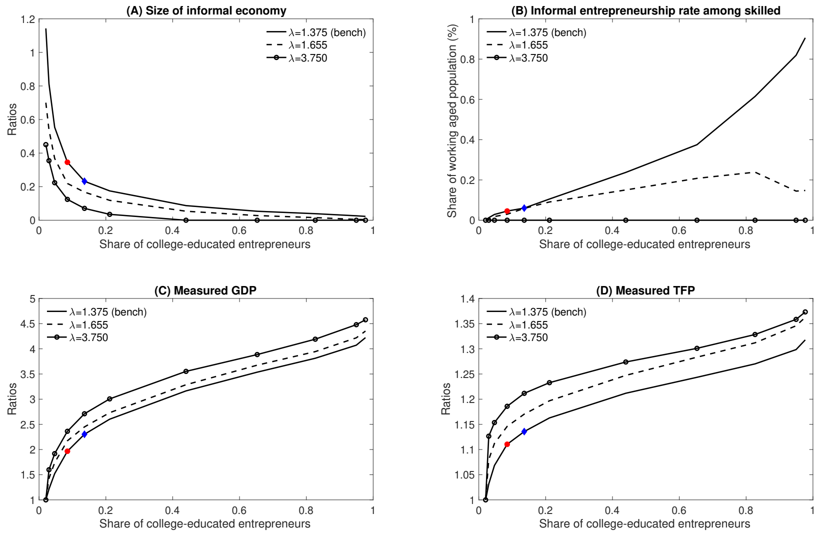


Notes: The figure reports steady-state outcomes obtained by varying $\bar{\kappa}$ while holding all other parameters fixed. The x-axis measures the equilibrium share of college-educated entrepreneurs, which increases as $\bar{\kappa}$ declines. In Panel D, the informal entrepreneurship rate among unskilled individuals is measured on the left axis, while the corresponding rate among skilled individuals is measured on the right axis. A red circle denotes the calibrated benchmark economy, whereas a blue diamond denotes the counterfactual equilibrium with a population college rate of 12.5%.

ercise for alternative values of the collateral constraint parameter, λ. Figure 6 reports the results, showing that, as access to credit improves, the decline in firm informality associated with higher college attainment becomes substantially stronger. In particular, for sufficiently high values of λ, increasing the college rate drives the size of the informal economy to zero. The reason is that, once access to credit is sufficiently high, matched college-educated entrepreneurs no longer have an incentive to operate informally. The figure makes this mechanism explicit: under the highest value of λ considered, which implies a credit-to-GDP ratio of about 130 percent and thus corresponds to an economy with very limited financial frictions, the share of skilled individuals engaged in informal entrepreneurship is zero for all levels of college attainment (see Panel B). The results therefore confirm that human capital accumulation and access to credit are complementary forces: higher education expands the pool of high-ability potential entrepreneurs, while better credit access allows these entrepreneurs to operate at the scale needed to make immediate formalization optimal.

Overall, the comparative-statics exercises in this section also generalize the aggregate results of Section 5.2 by showing that improvements in educational attainment generate gains in official GDP and measured TFP that become larger as the college rate increases (see Figure 5). Panels C and D of Figure 6 further show that these effects are amplified when higher educational attainment is accompanied by better access to credit, reinforcing the complementarity between human capital accumulation and financial development predicted by our model.

Figure 6: Interaction between human capital and credit



Notes: The figure repeats the comparative-statics exercise over $\bar{\kappa}$ for three values of the collateral constraint parameter, λ. The benchmark value is $\lambda = 1.375$; the other two values are chosen so that, holding all other parameters fixed, the model generates credit-to-GDP ratios of 0.70 and 1.30, respectively. In Panels C and D, measured GDP and measured TFP are normalized separately for each value of λ, using the corresponding economy with the lowest share of college-educated entrepreneurs as the baseline.

6 Empirical evidence

In this section, we test model implications using Brazilian data. In particular, we are interested in understanding whether the implications of human capital accumulation on firm informality predicted by the model are consistent with the data. Brazil has seen significant advancements in tertiary education completion rates throughout the 2000s, presenting a compelling case study. Specifically, the proportion of individuals who have completed at least short-cycle tertiary education nearly doubled from 2004 to 2023, while the gross enrollment rate in tertiary education more than tripled between 1999 and 2022 (UNESCO Institute for Statistics (UIS), 2024).[35]

This growth in enrollment can be primarily attributed to the expansion of private for-profit higher education institutions following the 1996 education law, which simplified entry for higher education providers (Cox, 2024). Specifically, private institutions were allowed to operate on a for-profit basis, and the requirements for establishing new colleges were streamlined. Furthermore, as highlighted in Cox (2024), this expansion—driven largely by private institutions—lowered barriers to college attainment in two key ways: by reducing commuting distances and easing capacity constraints. In light of this, the reform provides

[35]School enrollment in tertiary education is defined as a ratio of total enrollment, regardless of age, to the population of the age group that officially corresponds to the level of education. The population of the official age for tertiary education is estimated to be the 5-year age group immediately following upper secondary education. If the official entrance age to upper secondary is 15 years and the duration is 3 years, then a is the age group 18-22 years.

a natural setting to test whether the model's predictions align with observed patterns of informality in the data.

We structure our empirical analysis using a two-step approach. First, we leverage census data to document a negative correlation between human capital and informality at the municipal level. Furthermore, we show that municipalities with larger increases in college attainment rates have experienced larger reductions in informal entrepreneurship, in line with the predictions of our model. Next, we move beyond correlational analysis to establish a causal link between college attainment and firm informality by employing a difference-in-difference approach. Specifically, following Cox (2024), we exploit the 1996 educational reform, which lowered barriers to college attainment, to identify its impact on informal entrepreneurship. One of the features of this reform is that the expansion of higher education in Brazil was primarily driven by for-profit institutions, making college entry largely dependent on market forces rather than centralized planning. This setting allows us to classify areas into high- and low-intensity exposure to higher education, based on the potential demand for college education, following the approach of Duflo (2001). Leveraging variation in both age and geographical exposure to excess college entry, we find that individuals in areas with relatively higher college entry rates were less likely to engage in informal entrepreneurship. These results further support the implications of our model.

6.1 Reduced form evidence

In the model, there is a negative relationship between college attainment and firm informality. To test empirical plausibility of this relationship, we utilize data from the Brazilian 2000 and 2010 censuses, focusing on the period prior to the effects of the expansion of higher education in 1996 and after.[36] Our findings confirm that this relationship holds at both the municipality and commuting zone levels (*Áreas Mínimas Comparáveis*).[37] In particular, this level of granularity allows us to capture how variations in local educational expansion correspond to differences in informal entrepreneurship.

We compute municipality and commuting zone-level rates of college attainment and informal entrepreneurship among the working-age urban population, and run the following cross-sectional regressions for 2000 and 2010:

$$Infe_{m,t} = \alpha + \beta_1 CollegeShare_{m,t} + \beta_2 log(GDP_{m,t}) + \beta_3 log(Credit_{m,t}) + \epsilon_{m,t} \tag{10}$$

where $CollegeShare_{m,t}$ denotes rates of college attainment in a municipality (commuting zone) m for year t, $log(GDP_{m,t})$ and $log(Credit_{m,t})$ represent the logarithm of GDP per capita and credit per capita, respectively, for a municipality (commuting zone) m in year t. $Infe_{m,t}$ denotes the rates of informal entrepreneurship in a municipality (commuting zone) m in year t.

[36]Chamber of higher education of Brazilian Ministry of Education establishes guidelines for completion of higher education degree. In Brazil college attainment requires 4-5 years. The 2000 census was chosen as a reference year because it is sufficiently close to the time of the higher education reform but still precedes the period when the reform's effects would have fully materialized.

[37]As noted by Cox (2024) and Dix-Carneiro and Kovak (2017), commuting zones represent stable and comparable labor markets over time, providing reassurance of the relationship across different geographic units.

Table 8: Negative correlation between rates of informal entrepreneurship and college-attainment (cross-municipality)

	Municipality Level		CZ Level	
	(1)	(2)	(3)	(4)
College Rates	-0.326***	-0.218***	-0.271**	-0.326***
	(0.0475)	(0.0340)	(0.113)	(0.0956)
log(Creditpc)	0.000812	-0.00304***	-0.00772*	-0.00555
	(0.00116)	(0.00116)	(0.00415)	(0.00433)
log(GDPpc)	-0.0766***	-0.0549***	-0.0748***	-0.0434***
	(0.00292)	(0.00222)	(0.00572)	(0.00474)
Constant	0.387***	0.352***	0.458***	0.362***
	(0.00771)	(0.00795)	(0.0232)	(0.0269)
Agg. level	Municipality	Municipality	CZ	CZ
Year	2000	2010	2000	2010
N	3223	3426	470	469
F	372.0	420.7	151.2	137.7
R^2	0.373	0.335	0.563	0.466

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are reported in parentheses.

Estimation results are presented in Table 8. As observed, there is a strong statistically significant negative relationship between college attainment and informality. In particular, our results indicate that at the municipality (commuting zone) level, after controlling for economic development through GDP per capita and credit availability, a 1 percentage point increase in the college share is associated with a 0.326 (0.271) percentage point decrease in the informal entrepreneurship rate in 2000 and a 0.218 (0.326) percentage point decrease in 2010. Additionally, the negative coefficient on credit per capita suggests that greater credit availability is linked to lower rates of informal entrepreneurship. Overall, these findings support a statistically significant negative relationship between human capital and firm informality across Brazilian municipalities and commuting zones, both before and after the effects of higher educational policies materialized.

We complement our empirical analysis with a further test of the model predictions. According to the model, higher rates of college attainment should lead to a substantial reduction in the size of the informal sector. Consequently, we examine whether municipalities with larger increases in college attainment also experienced larger reductions in informality rates, controlling for changes in economic development and the initial levels of college attainment and informality. To this end, we estimate the following regression model at the municipality (commuting zone) level m:

$$\begin{aligned}\Delta Infe_m = \alpha + \beta_1 Y_{m,2000} + \beta_2 College_{m,2000} + \beta_3 logCreditpc_{m,2000} + \beta_4 logGDPpc_{m,2000} + \\ \beta_5 \Delta College_m + \beta_6 \Delta logCreditpc_m + \beta_7 \Delta logGDPpc_m + \epsilon_m \end{aligned} \quad (11)$$

where $\Delta Infe_m$ represents the change in informal entrepreneurship rates between 2000 and 2010, and $Y_{m,2000}$ denotes the initial rate in municipality (CZ) m in 2000. Results are presented in Table 9. The first row indicates that, after controlling for initial levels of college attainment and GDP per capita, municipalities that experienced larger increases in college attainment between 2000 and 2010 also saw greater reductions in informal entrepreneurship rates. A similar pattern emerges at the commuting zone level. The observed negative

Table 9: Regression coefficients for Equation 11

	(1)		(2)	
Diff College	-0.115***	(0.0387)	-0.194**	(0.0900)
Diff LogCreditpc	0.00203	(0.00147)	0.00918**	(0.00379)
Diff LogGDPpc	-0.00906***	(0.00238)	-0.000167	(0.00547)
Constant	0.0989***	(0.00804)	0.0524***	(0.0201)
Agg Level	Municipality		CZ	
College2000	Yes		Yes	
LogGDPpc2000	Yes		Yes	
LogCreditpc2000	Yes		Yes	
N	3012		466	
F	150.8		47.41	
R^2	0.416		0.579	

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are reported in parentheses.

and statistically significant relationship corroborates predictions of the model. While this analysis remains correlational, we take this evidence as additional support for the role of human capital in shaping firm informality.

6.2 Difference-in-difference approach

This section explores the link between human capital and firm informality using a quasi-experimental framework. While previous empirical analyses have primarily been correlational, Brazilian data and institutional context provide an opportunity to evaluate the potential causal impact of college attainment on the incentives to operate an informal firm. However, despite the considerable interest, evidence of the causal influence of tertiary education—unlike that of elementary education—on economic development remains limited.[38] A common concern in the literature studying the effects of educational attainment on socio-economic outcomes is the presence of confounding factors (see, e.g., Duflo, 2001; Osili and Long, 2008; Akresh *et al.*, 2022). To address this challenge, studies often rely on quasi-experimental approaches such as regression discontinuity, difference-in-differences, and instrumental variables (Khanna, 2023; Machado *et al.*, 2023).

In line with this, we exploit the quasi-natural experiment created by the 1996 educational reform to assess its impact on the probability of running an informal firm. Our identification strategy leverages variation in exposure to the reform based on individuals' age and place of residence. Since college attendance in Brazil typically occurs between ages 16 and 30, we define the treatment group as individuals aged 16–30, who were exposed to the reform during their prime college years, and the control group as individuals aged 31–45, who were older than the typical college-going age at the time of the reform.

Beyond age-based variation, we also account for geographical differences in treatment intensity. We define place of residence at the level of commuting zones, which serve as relevant geographic units for local labor and education markets.[39] As noted earlier, the expansion of higher education was primarily driven by market demand. Leveraging this

[38] Notable exceptions include Cox (2024) and Machado *et al.* (2023).

[39] Brazil has experienced an increase in the number of municipalities since 1991. Because we rely on higher education census data from 1995, it is essential to define a stable geographical unit that remains comparable over time. To achieve this, we follow the methodology in Ehrl (2016), which constructs consistent commuting zones. This approach balances geographic granularity with migration dynamics, ensuring that migration flows across commuting zones are minimized within the chosen classification.

feature, we measure the intensity of college entry into commuting zones based on the potential demand for higher education. Using higher education census data from 1995 and 2005, we classify commuting zones into high- and low-intensity treatment areas, following the approach of Duflo (2001). Specifically, we hypothesize that individuals in commuting zones with disproportionately high college entry rates relative to the potential pool of college students are more likely to pursue and complete higher education, thereby reducing their likelihood of engaging in informal entrepreneurship.

6.2.1 Data

We use Brazilian decennial censuses from 2000 and 2010, which provide detailed information on education, migration, demographics, and occupational characteristics (Instituto Brasileiro de Geografia e Estatística (IBGE), 1991, 2000, 2010). To evaluate the outcomes of the educational reform, we use individual-level data on socio-economic and demographic characteristics from the 2010 census. To classify commuting zones into high- and low-intensity treatment groups, we rely on higher education census data from 1995 and 2005 (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), 1995,2005), which allows us to determine the total number of colleges in each commuting zone. To estimate the potential demand for higher education prior to the reform, we use the 1991 census to approximate the applicant pool at the commuting zone level.[40]

6.2.2 Empirical Specification

To provide an initial benchmark, we begin by estimating an OLS regression at the individual level:

$$Infe_{ij} = \alpha + \boldsymbol{X'_{ij}\theta} + \beta College_{ij} + \phi_j + \epsilon_{ij} \tag{12}$$

where $Infe_{ij}$ is a binary indicator for informal entrepreneurship, equal to 1 if individual i in commuting zone j is an informal entrepreneur and 0 if the individual is a worker or a formal entrepreneur. The key explanatory variable, $College_{ij}$, equals 1 if the individual has completed a college degree. This estimation controls for commuting zone fixed effects (ϕ_j) and a set of individual characteristics, including gender, race, marital status, and whether the individual has always resided in the commuting zone. As shown in Table 10, the OLS estimates suggest that individuals with college education are significantly less likely to engage in informal entrepreneurship. However, these results may suffer from endogeneity issues, including self-selection into college and unobserved factors that influence both educational attainment and occupational choice.

To establish a causal relationship between college attainment and firm informality, we employ a difference-in-differences approach, estimating the following equation:

$$Infe_{ij} = \alpha + \boldsymbol{X'_{ij}\theta} + \beta T_i + \gamma(T_i \times HI_j) + \phi_j + \epsilon_{ij} \tag{13}$$

[40]The potential applicant pool is defined as the weighted mass of individuals aged 20-29 in the 1991 census who had completed high school but did not hold a college degree. As a robustness check, we construct an alternative applicant pool using individuals aged 19–33 in the 2000 Census; the corresponding results are reported in Appendix C.2 Table C4.

Table 10: OLS Regression Results for Equation 12.

	(1) Infe	(2) Infe	(3) Infe	(4) Infe	(5) Infe
College Degree	-0.061***	-0.064***	-0.061***	-0.059***	-0.054***
	(0.001)	(0.005)	(0.001)	(0.001)	(0.001)
Single=1					-0.010***
					(0.001)
White=1					0.013***
					(0.001)
Always lived CZ=1					-0.004***
					(0.001)
Male=1					0.054***
					(0.001)
FE	CZ	UF	CZ	CZ-Age	CZ-Age
Cluster	Robust	CZ	CZ-Age	CZ-Age	CZ-Age
Observations	1,857,453	1,857,453	1,857,453	1,857,453	1,857,452

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are reported in parentheses. Population weights are applied. UF corresponds to state fixed effects.

where $Infe_{ij}$ is as defined before. T_i is an indicator for whether an individual belonged to the younger cohort, taking a value of 1 for those aged 16–30 in the reform year, 1997, and 0 for those aged 31–45, who represent the older cohort.[41] HI_j is an indicator of whether a commuting zone j experienced disproportionate entry of colleges relative to the pool of potential college applicants. This estimation controls for commuting zone fixed effects, ϕ_j, and observable individual characteristics, as in the previous regression.

We classify commuting zones (CZs) into high- and low-intensity treatment groups following Duflo (2001), using the residuals from the following regression:

$$C_j = \alpha_0 + \beta StudentPool_j + \epsilon_j \tag{14}$$

where C_j represents a change in the number of colleges in CZ j between 1995 and 2005.[42] $StudentPool_j$ is measured as a mass of individuals aged 20-29 years in the year 1991 within commuting zone j who have obtained a high school diploma but have not obtained a college degree. We assign $HI_j = 1$ to commuting zones that experienced an excessive entry of colleges relative to the potential pool of applicants. Specifically, commuting zones with positive residuals from the regression, after controlling for the potential applicant pool, are classified as high-intensity treatment areas.[43]

Table 11 reports the estimation results. The first row presents the coefficient on exposure, T_i, which captures the conditional difference in informal entrepreneurship between younger and older individuals in low-intensity commuting zones. In column (4), this coefficient is negative and statistically significant, suggesting that, all else equal, the younger cohort was 3.9 percentage points less likely to engage in informal entrepreneurship. Our main coefficient

[41] 1997 is taken as the reference year since the educational law was passed in 1996, with additional decrees enacted in 1997 that established the foundation for the reform (Cox, 2024).

[42] As robustness, we use an alternative measure of intensity, defining C_j as an absolute number of colleges in 2005. Results are reported in Table C3 in the Online Technical Appendix C.2. Moreover, instead of a binary indicator we provide results using a continuous measure of intensity, defined as the standardized residuals from Equation 14. Results are reported in Appendix C.2 Table C5.

[43] Figure C.1 in the Online Technical Appendix C.3 plots the spatial distribution of high-intensity commuting zones.

Table 11: Regression coefficients for Equation 13.

	(1) Infe	(2) Infe	(3) Infe	(4) Infe	(5) Infe
Born 1967-1981=1	-0.042***	-0.033***	-0.042***	-0.039***	
	(0.001)	(0.002)	(0.002)	(0.002)	
HI × Born 1967-1981=1	-0.008***	-0.023***	-0.008***	-0.007***	-0.008***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)
Single=1				-0.009***	-0.007***
				(0.001)	(0.001)
Always lived CZ=1				-0.006***	-0.004***
				(0.001)	(0.001)
Male=1				0.060***	0.060***
				(0.001)	(0.001)
White=1				0.003***	0.003***
				(0.001)	(0.001)
FE	CZ	UF	CZ	CZ	CZ-Age
Cluster	Robust	CZ-Age	CZ-Age	CZ-Age	CZ-Age
Young low-intensity mean	0.138	0.138	0.138	0.138	0.138
Observations	3,229,366	3,229,366	3,229,366	3,229,363	3,229,363

Notes: Significance levels: * 0.10 ** 0.05 *** 0.01. Standard errors are reported in parentheses. Population weights are applied. High-intensity treatment is defined as the change in number of colleges between 1995 and 2005. UF corresponds to state fixed effects.

of interest—the interaction between high-intensity exposure and the treated cohort—is also negative and statistically significant across all specifications. This finding indicates that the decline in informal entrepreneurship among younger individuals was more pronounced in high-intensity commuting zones than in low-intensity commuting zones. In column (4), the interaction coefficient implies that *relative to the treated cohort* young individuals in high-intensity areas were 17.95% (0.007/0.039) less likely to engage in informal entrepreneurship compared to their counterparts in low-intensity areas. Our preferred specification is reported in column (5), which adds commuting-zone and age fixed effects. The interaction coefficient is -0.008, equivalent to 5.8% of the informal-entrepreneurship rate among younger individuals in low-intensity areas (0.008/0.138).

To provide support for the validity of our method, we assess the plausibility of the parallel trends assumption by conducting placebo tests using fake treatment groups and alternative reform dates. The results, which provide evidence in favor of our identification strategy, are presented in the Online Technical Appendix C.2. To summarize, our findings demonstrate that the reform reducing barriers to college attainment led to a decline in informal entrepreneurship, particularly in commuting zones with excessive college entry. Coupled with the results of the reduced form estimations, these findings lend further support to the main predictions of our model.

7 Conclusions

This paper builds a life-cycle general equilibrium model to investigate how entrepreneurial human capital, shaped by endogenous education decisions and subject to mismatch risk, affects firm informality in developing economies. By embedding capital-skill complementarity and financial frictions into a model of occupational choice under imperfect tax enforcement, we show that access to college education plays a central role in determining the size and composition of the informal sector.

Beyond enforcement and credit constraints, our results underscore that the endogenous distribution of entrepreneurial human capital plays a first-order role in shaping occupational choices and firm formalization. Calibrated to Brazilian data in 2003, the model shows that college-educated individuals are more likely to become formal entrepreneurs running larger and more productive firms, while less-educated individuals are disproportionately represented among informal enterprises. The key mechanism is dynamic: entrepreneurial human capital leads firms to enter at larger scale and grow faster over the life cycle. As firms expand, they face higher detection risk and greater demand for credit, both of which reduce the relative attractiveness of informality.

Counterfactual analysis shows that raising college attainment to the level Brazil reached in 2012 would reduce the size of the informal sector by 11.3 percentage points, raise official GDP by 7.7%, and increase productivity and fiscal revenues, with roughly half of these gains attributable to entrepreneurial human capital rather than the wage-skill premium alone. At the same time, the model points to an important qualification: financial frictions can still make informal entrepreneurship optimal for newly skilled entrepreneurs, so informality among this group rises even as aggregate informality falls. Educational improvements alone are therefore insufficient to eliminate firm informality. Full formalization requires improvements in access to education and credit markets to occur jointly.

We provide empirical support for the mechanism using Brazil's 1996 Higher Education Reform, which reduced barriers to entry for private higher education institutions and generated differential expansions in college access across cohorts and regions. Using a difference-in-differences framework, we exploit variation in exposure to the reform, proxied by excess college entry, and show that individuals in more exposed areas were significantly less likely to become informal entrepreneurs. Together with the cross-sectional evidence linking college attainment and informality across Brazilian municipalities, these results support the model's prediction that expanding access to higher education can reduce informal entrepreneurship.

References

Akresh, R., Halim, D. and Kleemans, M. (2022). 'Long-term and intergenerational effects of education: Evidence from school construction in indonesia', *The Economic Journal*, vol. 133(650), pp. 582–612.

Allen, J., Nataraj, S. and Schipper, T.C. (2018). 'Strict duality and overlapping productivity distributions between formal and informal firms', *Journal of Development Economics*, vol. 135, pp. 534–554.

Allub, L. and Erosa, A. (2019). 'Financial frictions, occupational choice and economic inequality', *Journal of Monetary Economics*, vol. 107(C), pp. 63–76.

Allub, L., Gomes, P. and Kuehn, Z. (2023). 'Human Capital and Financial Development: Firm-Level Interactions and Macroeconomic Implications', *The Economic Journal*, vol. 134(659), pp. 934–958.

Amaral, P.S. and Quintin, E. (2006). 'A competitive model of the informal sector', *Journal of Monetary Economics*, vol. 53(7), pp. 1541–1553.

Antunes, A. and Cavalcanti, T. (2007). 'Start up costs, limited enforcement, and the hidden economy', *European Economic Review*, vol. 51(1), pp. 203–224.

Antunes, A., Cavalcanti, T. and Villamil, A. (2015). 'The effects of credit subsidies on development', *Economic Theory*, vol. 58(1), pp. 1–30.

Autor, D.H. (2014). 'Skills, education, and the rise of earnings inequality among the "other 99 percent"', *Science*, vol. 344(6186), pp. 843–851.

Belzil, C. and Hansen, J. (2002). 'Unobserved ability and the return to schooling', *Econometrica*, vol. 70(5), pp. 2075–2091.

Berniell, L. (2021). 'Occupational choice and investments in human capital in informal economies', *The B.E. Journal of Macroeconomics*, vol. 21(2), pp. 399–423.

Bobba, M., Flabbi, L. and Levy, S. (2022). 'Labor market search, informality, and schooling investments', *International Economic Review*, vol. 63(1), pp. 211–259.

Bobba, M., Flabbi, L., Levy, S. and Tejada, M. (2021). 'Labor market search, informality, and on-the-job human capital accumulation', *Journal of Econometrics*, vol. 223(2), pp. 433–453.

Buera, F. (2009). 'A dynamic model of entrepreneurship with borrowing constraints: Theory and evidence', *Annals of Finance*, vol. 5(3), pp. 443–464.

Buera, F., Kaboski, J.P. and Shin, Y. (2011). 'Finance and development: A tale of two sectors', *American Economic Review*, vol. 101(5), pp. 1964–2002.

Buera, F. and Shin, Y. (2013). 'Financial frictions and the persistence of history: A quantitative exploration', *Journal of Political Economy*, vol. 121(2), pp. 221–272.

Castro, J.F., Ortega, L., Yamada, G. and Mata, D. (2024). 'Overeducation and overskilling in latin america: Evidence from piaac', *Comparative Education*, vol. 60(2), pp. 296–314.

Cavalcanti, T. and Santos, M. (2020). '(Mis)Allocation effects of an overpaid public sector', C.E.P.R. Discussion Papers.

Ciccone, A. and Papaioannou, E. (2009). 'Human capital, the structure of production, and growth', *The review of economics and statistics*, vol. 91(1), pp. 66–82.

Cox, A. (2024). 'From classroom to prosperity: Fostering development through higher education', Working Paper.

Cunha, F. and Heckman, J. (2007). 'The technology of skill formation', *American Economic Review*, vol. 97(2), p. 31–47, doi:10.1257/aer.97.2.31.

de Paula, A. and Scheinkman, J.A. (2007). 'The Informal Sector, Third Version', Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.

de Paula, A. and Scheinkman, J.A. (2010). 'Value-added taxes, chain effects, and informality', *American Economic Journal: Macroeconomics*, vol. 2(4), pp. 195–221.

de Paula, A. and Scheinkman, J.A. (2011). 'The Informal Sector: An Equilibrium Model And Some Empirical Evidence From Brazil', *Review of Income and Wealth*, vol. 57, pp. 8–26.

D'Erasmo, P. (2013). 'Access to credit and the size of the formal sector in Brazil', Inter-American Development Bank, Research Department.

D'Erasmo, P.N. and Boedo, H.J.M. (2012). 'Financial structure, informality and development', *Journal of Monetary Economics*, vol. 59(3), pp. 286–302.

Di Nola, A., Kocharkov, G., Scholl, A. and Tkhir, A.M. (2021). 'The aggregate consequences of tax evasion', *Review of Economic Dynamics*, vol. 40, pp. 198–227.

Diaz, J.J., Chacaltana, J., Rigolini, J. and Ruiz, C. (2018). 'Pathways to formalization: Going beyond the formality dichotomy', Institute of Labor Economics (IZA).

Dix-Carneiro, R. and Kovak, B.K. (2017). 'Trade liberalization and regional dynamics', *American Economic Review*, vol. 107(10), p. 2908–46.

Duflo, E. (2001). 'Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment', *American Economic Review*, vol. 91(4), p. 795–813.

Ehrl, P. (2016). 'Minimum comparable areas for the period 1872–2010: An aggregation of brazilian municipalities', *Revista Brasileira de Economia*, vol. 71(4), pp. 215–229.

Erosa, A. (2001). 'Financial intermediation and occupational choice in development', *Review of Economic Dynamics*, vol. 4(2), pp. 303–334.

Erosa, A., Fuster, L. and Martinez, T.R. (2023). 'Public financing with financial frictions and underground economy', *Journal of Monetary Economics*, vol. 135, pp. 20–36.

Fajardo, J., Ornelas, J.R.H. and Farias, A.R.d. (2012). 'Estimating risk aversion, risk-neutral and real-world densities using Brazilian Real currency options', *Economia Aplicada*, vol. 16, pp. 567 – 577.

Fernández, M. and Messina, J. (2018). 'Skill premium, labor supply, and changes in the structure of wages in latin america', *Journal of Development Economics*, vol. 135, pp. 555–573.

Fernández-Bastidas, R. (2023). 'Entrepreneurship and tax evasion', *Economic Modelling*, vol. 128, p. 106488.

Ferreira, P. and Veloso, F. (2013). 'O desenvolvimento econômico brasileiro no pós-guerra', in (F. Veloso, P. C. Ferreira, F. Giambiagi and S. Pessôa, eds.), *Desenvolvimento Econômico: Uma Perspectiva Brasileira*, pp. 129–165, Rio de Janeiro: Elsevier.

Firpo, S. and Portella, A. (2019). 'Decline in wage inequality in brazil : A survey', The World Bank, doi:None.

Fonseca, J. and Van Doornik, B. (2022). 'Financial development and labor market outcomes: Evidence from brazil', *Journal of Financial Economics*, vol. 143(1), pp. 550–568.

Franjo, L., Pouokam, N. and Turino, F. (2022). 'Financial Frictions and Firm Informality: A General Equilibrium Perspective', *The Economic Journal*, vol. 132(645), pp. 1790–1823.

Haanwinckel, D. and Soares, R.R. (2021). 'Workforce Composition, Productivity, and Labour Regulations in a Compensating Differentials Theory of Informality', *The Review of Economic Studies*, vol. 88(6), pp. 2970–3010.

Heathcote, J., Storesletten, K. and Violante, G.L. (2010). 'The macroeconomic implications of rising wage inequality in the united states', *Journal of political economy*, vol. 118(4), pp. 681–722.

Heathcote, J., Storesletten, K. and Violante, G.L. (2017). 'Optimal tax progressivity: An analytical framework', *The Quarterly Journal of Economics*, vol. 132(4), pp. 1693–1754.

Instituto Brasileiro de Geografia e Estatística (IBGE) (1991, 2000, 2010). 'Censo demográfico', `https://www.ibge.gov.br/en/statistics/social/labor/22836-2022-census-3.html`, accessed: [2023].

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) (1995,2005). 'Microdados do censo da educação superior', Accessed: [2024].

Jung, J. and Tran, C. (2012). 'The extension of social security coverage in developing countries', *Journal of Development Economics*, vol. 99(2), pp. 439 – 458.

Khanna, G. (2023). 'Large-scale education reform in general equilibrium: Regression discontinuity evidence from india', *Journal of Political Economy*, vol. 131(2), pp. 549–591.

Krusell, P., Ohanian, L.E., Ríos-Rull, J.V. and Violante, G.L. (2000). 'Capital-skill complementarity and inequality: A macroeconomic analysis', *Econometrica*, vol. 68(5), pp. 1029–1053.

La Porta, R. and Shleifer, A. (2008). ‘The Unofficial Economy and Economic Development’, *Brookings Papers on Economic Activity*, vol. 39(2 (Fall)), pp. 275–363.

La Porta, R. and Shleifer, A. (2014). ‘Informality and development’, *Journal of Economic Perspectives*, vol. 28(3), pp. 109–126.

Leal Ordoñez, J.C. (2014). ‘Tax collection, the informal sector, and productivity’, *Review of Economic Dynamics*, vol. 17(2), pp. 262–286.

Lucas, R.E. (1978). ‘On the size distribution of business firms’, *Bell Journal of Economics*, vol. 9(2), pp. 508–523.

Machado, C., Reyes, G. and Riehl, E. (2023). ‘The direct and spillover effects of large-scale affirmative action at an elite brazilian university’, *Journal of Labor Economics*, vol. 0(ja), p. null.

Marioni, L.d.S. (2021). ‘Overeducation in the labour market: Evidence from brazil’, *Education Economics*, vol. 29(1), pp. 53–72.

Medina, L. and Schneider, F. (2018). ‘Shadow economies around the world: What did we learn over the last 20 years?’, International Monetary Fund.

Meghir, C., Narita, R. and Robin, J.M. (2015). ‘Wages and informality in developing countries’, *American Economic Review*, vol. 105(4), pp. 1509–1546.

Morazzoni, M. (2021). ‘Student Debt and Entrepreneurship in the US’, Available at SSRN 4440926.

Osili, U.O. and Long, B.T. (2008). ‘Does female schooling reduce fertility? evidence from nigeria’, *Journal of Development Economics*, vol. 87(1), pp. 57–75.

Perry, G.E., Maloney, W.F., Arias, O.S., Fajnzylber, P., Mason, A.D. and Saavedra-Chanduvi, J. (2007). *Informality: Exit and exclusion*, no. 6730 in World Bank Publications, The World Bank.

Quadrini, V. (2000). ‘Entrepreneurship, saving and social mobility’, *Review of Economic Dynamics*, vol. 3(1), pp. 1–40.

Queiró, F. (2021). ‘Entrepreneurial Human Capital and Firm Dynamics’, *The Review of Economic Studies*, vol. 89(4), pp. 2061–2100.

Quintin, E. (2008). ‘Contract enforcement and the size of the informal economy’, *Economic Theory*, vol. 37(3), pp. 395–416.

Soares, R.R. and Haanwinckel, D. (2017). ‘Fighting employment informality with schooling’, *IZA World of Labor*, pp. 394–394.

Ulyssea, G. (2018). ‘Firms, informality, and development: Theory and evidence from Brazil’, *American Economic Review*, vol. 108(8), pp. 2015–47.

UNESCO Institute for Statistics (UIS) (2024). ‘UIS.Stat Bulk Data Download Service’, Indicators: SE.TER.CUAT.ST.ZS, SE.TER.ENRR. Accessed September 30, 2024.

Van Reenen, J. (2011). ‘Wage inequality, technology and trade: 21st century evidence’, *Labour economics*, vol. 18(6), pp. 730–741.