# Human Capital and Growth: The Role of High-Skill Labor Concentration

Julio Brandao-Roll
LSE*

**Abstract**

This paper raises and tests the hypothesis that the effects of human capital on economic growth depend crucially on the concentration of high-skill labor across firms. Importantly and surprisingly, an increase in human capital supply can actually lower growth if skill concentration across firms is high enough. Intuitively, large firms have limited financial incentives to innovate because they dominate the market and incur the risk of self-cannibalization when innovating; therefore, when increased skill supply primarily benefits these firms, the equilibrium growth impacts can be negative. I investigate this hypothesis in Brazil, establishing three results. First, in a difference-in-differences design across municipalities, I estimate that new colleges had a positive impact on local economic growth in municipalities with lower concentration of high-skill labor, but a negative effect in municipalities with higher skill concentration. Second, I isolate the causal effect of changes in local high-skill labor concentration on local growth using a shift-share design, leveraging loan shocks to firms. Third, I develop and estimate an endogenous growth model, which quantitatively matches the preceding results and which I use to assess policy counterfactuals. These results help explain why several middle-income countries, including Brazil, have experienced a slowdown in growth despite a fast increase in high-skill supply over the past decades.

# 1 Introduction

Why do increases in human capital supply, particularly in developing countries, not always lead to higher growth rates? A longstanding macroeconomic literature associates improvements in human capital with higher economic growth, either through better labor productivity (Mankiw, Romer and Weil, 1992) or innovation (Romer, 1990). This has led to a push in the last decades by national governments and international organizations for an accumulation of skills, especially in middle-income countries, under the assumption that low-growth countries lack human capital. The empirical evidence on this positive associ-

*Contact: j.brandao-roll@lse.ac.uk. Web: https://www.juliobrandaoroll.com/. I am particularly indebted to Xavier Jaravel and Maarten De Ridder for extensive advice and comments. I also thank Philippe Aghion, Oriana Bandiera, Scott Behmer, Felipe Bordini, Matthias Doepke, Arnaud Dyèvre, Jonathon Hazell, Alan Manning, Guy Michaels, Ben Moll, Joana Naritomi, Elio Nimier-David, Jane Olmstead-Rumsey, Derek Pillay, Steve Pischke, Ricardo Reis, Daniel Sturm, Gabriel Ulyssea, participants in the LSE Labour and Public Economics WIP Seminars, the LSE Junior Macro Workshop, and at the 2024 Warwick Economics PhD Conference for comments and suggestions.

ation, however, has produced mixed results, mainly due to the data limitations and identification issues.[1] For instance, average TFP growth has been lackluster in middle-income countries since the mid-2000s even though high-skill labor supply soared, as shown in Figure A1 in the Appendix. This begs the question of whether there is something about skill demand, or lack thereof, that could explain the absence of higher economic growth from more skill supply.

To tackle this question, I propose a new channel that links high-skill supply and growth via high-skill concentration at large firms. This new channel works in two steps. First, an increase in skill supply raises local high-skill concentration, defined here as the local share of high-skill people working at large firms over total local supply. This is because a larger supply of human capital benefits large firms more than small ones, which further increases their relative gap. Second, this increase in skill concentration lowers firms' incentives to innovate once the gap between large and small firms is large enough. As the market leader grows larger, the incremental profit from further improvements over their own product keeps declining due to lower incremental gains in market share, a discouragement effect also known as Arrow's replacement effect (Arrow, 1962). These two steps create an offsetting effect that can cancel out, and even overcome, the positive effect on growth that we would expect from a reduction in innovation costs as skill supply increases. This novel channel can, then, explain why successive past increases in high-skill supply, particularly in developing countries, did not lead to higher growth rates and can actually induce a growth slowdown if skill concentration is high enough.

To test this novel channel empirically, I start by showing that the relationship between high-skill supply and growth depends on local high-skill concentration. I do this using municipality-level data from Brazil on new college and university creation in a difference-in-differences design where I compare municipalities that received a new college with those that did not. The estimation of the causal effect relies on the assumption that the choice of where to open a new college is unrelated to local growth trends. I substantiate this identification assumption by showing evidence of no pre-trends on growth, employment flows, and proxies of college demand, and by showing robustness of results to changes in the control group. Results show that places where high-skill concentration was high before the arrival of a new college grow around 10% less in the long term than places with low skill concentration. This relative difference in growth rates is due to an initial increase in local growth at municipalities with lower skill concentration, which subsides in the long term, and a decline in long-term growth at places with higher skill concentration, a counterintuitive result. This is evidence that local skill concentration plays a key

[1] C.f. Bils and Klenow (2000), Durlauf, Johnson and Temple (2005), Pritchett (2006).

role in moderating the link between skill supply and growth. Results also show around a 6% decline in long term growth from new colleges across all municipalities.

I then investigate the exact mechanism underlying this new channel by breaking it into two steps. In the first step, I exploit the same difference-in-differences design as the one above to pin-down the causal effect of higher high-skill supply on local skill concentration. I show that college creation has led to a rise in high-skill concentration of almost 12% a decade after students started graduating, which implies that large firms are the ones hiring most of the new college graduates. The magnitude of the effect of skill supply on local skill concentration is quite significant. If we extend this result nationally, ignoring missing intercept problems for a moment, the aggregate increase in skill supply can explain almost half of the national increase in high-skill concentration in the same period, which went from around 42% to 67%, as shown in Figure A2 in the Appendix.

This surprising result can be understood through the following thought experiment: assume two firms of different sizes compete in the same market by innovating on their products with high-skill labor, though only the market leader produces anything since it has the best product variety. Both firms face a strategic incentive to improve their products: the leader aims to make its product harder to copy, while the follower wants to catch up with the leader and become dominant in the market. However, if each innovation improves a firm's marginal cost, then when the leader innovates it is able to extract higher profits from its market share. This profit incentive is exclusive to the leader as the follower does not produce. Hence, when we make innovation less costly by raising skill supply, though both firms want to innovate at a higher rate, the leading firm wants to innovate relatively more because it has an extra incentive to do so. This implies the larger firm will increase its relative share of high-skill hiring, which raises skill concentration.

In the second step, I present novel evidence that local high-skill labor concentration has a non-monotonic relationship with GDP growth. To identify the causal link between skill concentration and growth, we require random variation in local skill concentration. I achieve this through a shift-share instrumental variable (SSIV) design that leverages heterogeneous municipality exposure to national changes in the loan portfolio of the Brazilian Development Bank (BNDES). Importantly, identification relies on changes to loan amounts for different economic sectors being as-good-as-random, an assumption which I test through different falsification tests. As large (small) firms are as good as randomly allocated loans, local skill concentration rises (falls) as firms use such loans to hire skilled labor. Results show that at low levels of concentration, increasing skilled labor at large firms boosts local growth rates. This trend, however, reverses at higher concentration levels when the relationship is negative. These results, which visually characterize an inverted-U relationship, are evidence that incentives for firms to grow depend crucially on skill concentration.

We can, then, join these two steps to fully understand the high-skill concentration channel. As human capital supply increased from new colleges, large firms benefited relatively more than small firms which raised local skill concentration incrementally. Since skill concentration can lead to both higher or lower growth depending on its prior level, the final piece is to understand how Brazil moved along the curve between growth and skill concentration as the latter increased. We can extrapolate the reduced-form estimates to the aggregate economy, abstracting from missing intercept issues, and show that the rise in skill concentration due to the increase in human capital can potentially explain a decline in long-term growth rates of around 18.3% of average growth between 1999 and 2010.

Importantly, this is not the net effect of skill supply on growth but the partial effect through the skill concentration channel. For this calculation, we are assuming that the only effect of increasing skill supply is to raise skill concentration. This is useful as it highlights that the magnitude of the effect of the skill concentration channel can be quite significant, particularly in a country like Brazil where skill concentration was high. As the empirical estimate for the net effect of skill supply on growth, from the difference-in-differences estimation, is around 6%, we can conclude that the skill concentration channel can more than offset the positive effect of skill supply on growth. Hence, we can link the increase in human capital supply to a slowdown in growth in Brazil. This high-skill concentration channel is also likely to be applicable to other developing countries which saw a large boost to college enrollment without experiencing higher growth rates.

I rationalize these empirical findings with a step-by-step innovation model with firm strategic interaction and high-skill labor search. As in Aghion, Harris, Howitt and Vickers (2001), two firms compete in a duopoly through a quality ladder where a leading firm can be a number of steps ahead from a lagging one in terms of innovation. I then add two novel aspects to my model. First, I require both firms to search for the high-skill labor used as an R&D input as in the Diamond-Mortensen-Pissarides framework (Diamond, 1982, Mortensen, 1982, Pissarides, 1985). Adding search frictions allows firms to shed labor when incentives to hire are low, which is important for the mechanism, and allows the model to capture empirical trends in both high-skill unemployment and skill premium. Second, I make innovation catch up, also interpreted as R&D imitation, a function of high-skill labor.[2] This allows us to gauge how active R&D imitation changes with respect to higher skill supply, which I then link to innovation diffusion.

[2] An idea made explicit in Cohen and Levinthal (1989) although already mentioned earlier in analyzes of diffusion of specific technologies (Tilton, 1971, Evenson and Kislev, 1973)

I then bring the model to data and show that it is able to reproduce the non-monotonic empirical relationship between high-skill labor concentration and growth. I do so by linking both variables to the technological gap between leading and lagging firms. The intuition, which is similar to Aghion et al. (2001), is the following. Start with both firms at the same step in the quality ladder. When a firm innovates and moves ahead, R&D competition intensifies as the leading firm wants to defend its profit flow from the laggard's threat while the latter wants to catch up. As such, economic growth increases along with high-skill concentration as the leader has the additional profit incentive, which leads to relatively more hiring. As the technological gap keeps increasing, both firms face lower incentives to innovate. For the leader, the likelihood of the laggard ever catching up gets smaller and incremental profits from product improvements decline due to self-cannibalization, reducing the marginal benefit of innovating. As for the laggard, it faces disincentives to innovate as the gap between firms is large, implying a low likelihood of catching up, and any reduction of the gap results in a more intense competitive response from the market leader. As such, growth declines though skill concentration increases as the laggard's incentives to invest in R&D fall quicker than the leader's since the latter always earns some incremental profit from improvements in marginal cost. The corollary, then, is that skill concentration goes up while growth increases at first to then decline, leading to a non-monotonic relationship between both variables.

I can, then, use the model to study the effect of an increase in the supply of human capital. I do so for two scenarios. In the first one, I assume the increase in high-skill workers is inorganic, i.e. that it comes from an outside source. This implies a one-to-one increase in total population, which one would expect to mechanically increase economic growth. I then then use the model to decompose the effect of higher skill supply on growth into two opposing channels. On one hand, a larger skill supply boosts R&D output through a reduction in hiring costs, which has a positive effect on growth. On the other hand, since the leader benefits more from a higher supply of skills, the average gap between leader and follower increases. This has a negative effect on economic growth due to the stronger disincentives to innovate for both firms. Hence, whether growth increases or decreases depends on the strength of each one of these two channels. I show that growth stops increasing at a large enough high-skill supply level and that it even decreases in per-capita terms as the skill concentration channel gets stronger. In the second scenario, I assume the more realistic scenario where some of the increase in high-skill is a result of low-skill workers becoming high skill, for instance through education, making the former scarcer. Here, the model is able to capture additional empirical trends in Brazil: a decline in the skill premium and an increase in high-skill unemployment, both due to the leading firm becoming unwilling to absorb the extra high-skill supply.

Linking a growth slowdown to more high-skill supply is an important contribution for two reasons. First, it explains why educational policy might not produce higher economic growth and may even lead to a growth slowdown, a somewhat surprising result relative to the consensus.[3] What is key here is the role of skill concentration which can make education policy backfire because it ends up helping large firms grow even larger. Second, the possibility of a growth slowdown from higher skill supply is also not expected in several endogenous growth models. For instance, in a Romer-based model (Romer, 1990) an increase in the high-skill share, all else constant, similar to the one that happened in Brazil leads to an increase in the growth rate of approximately 63%.[4] More generally, the positive relationship between human capital and growth is a shared feature of endogenous growth models that follow Nelson and Phelps (1966) in associating higher education levels with more innovation.[5]

Finally, I assess two policy measures that a social planner could implement to counteract the skill concentration channel. In the first policy, I assume the planner can increase innovation diffusion from the leader to the laggard. I show that at high levels of high-skill supply the planner can increase the growth rate (1.6% vs. 1.45%) by reducing the technological gap between both firms, increasing competition and incentives to innovate. In the second policy, the planner is able to tax the local leading firm and use those funds to subsidize high-skill labor at the lagging firm. Helping laggards to "fight back" unlocks the expected growth boost from more high-skill supply. Results, then, highlight the important role of firm interaction in the link between human capital and growth, particularly in the form of skill concentration. Once we take this into account, the relevant policy lever in places where skill concentration is high becomes not only raising human capital but also improving competition policy and innovation catch-up by smaller firms.

*Related Literature*

My work relates to different strands of literature. In introducing the high-skill concentration channel, I shed new light in the relationship between human capital and economic growth. Previous studies have focused on two important channels that deliver a positive association between skills and growth. In the first one, a more educated workforce is

[3]From the World Bank: "Having a skilled workforce has been recognized as paramount to boosting competitiveness in an increasingly global and interdependent economic environment, fostering innovation and business creation and increasing productivity" (Roseth, Valerio, Gutiérrez et al., 2016).

[4]I show the Romer-based growth derivation in Section A.2 in the Appendix. As the Romer-based growth does not take into account shocks or labor utilization, it should be interpreted as a measure of *potential growth*.

[5]Another example is Aghion and Howitt (1992).

more productive (Becker, 1962, Lucas, 1988, Mankiw et al., 1992, Black and Lynch, 1996, Hanushek and Kimko, 2000) as we associate it with the quality of human capital. In the second one, improvements to human capital boost innovation, either by pushing the technological frontier or through higher adoption rates (Romer, 1990, Aghion and Howitt, 1992, Benhabib and Spiegel, 2005, Toivanen and Vaananen, 2016, Che and Zhang, 2017, Bianchi and Giorcelli, 2019, Biasi and Ma, 2022). My contribution is to propose a new channel where increasing high-skill supply raises high-skill concentration, which lowers growth. While in practice all three channels happen at the same time, I show that the high-skill concentration channel is useful in explaining the observed growth slowdown in Brazil and other developing countries where human capital soared. Empirically, by leveraging college creation I build on work showing the effects of colleges on local outcomes (Abramovsky, Harrison and Simpson, 2007, Toivanen and Vaananen, 2016, Azoulay, Graff Zivin, Li and Sampat, 2019, Valero and Van Reenen, 2019, Hausman, 2022, Nimier-David, 2023, Cox, 2024). Particularly, I follow Nimier-David (2023) in using college creation in an event study research design to identify the effect of a new college on both high-skill concentration and growth. I add to their results by showing heterogeneity with respect to the degree of local high-skill concentration on the effect of a new education establishment on local growth, which I rationalize in a model of step-by-step innovation.

In terms of both mechanism and model, I build on the large literature on endogenous growth, particularly on strategic interaction models (Aghion et al., 2001, Acemoglu and Akcigit, 2012, Liu, Mian and Sufi, 2022) to explain my findings. I link my results to the non-monotonicity induced by the "escape-competition" effect, where a market leader invests heavily in innovation to be further ahead of the competition, and the "lazy-monopolist" effect, where the leader stops investing when it is too far ahead. My contribution lies in adding high-skill labor demand and search to the step-by-step model which not only introduces the role of skill concentration but also extends results to the skill premium and high-skill unemployment. This paper also extends two previous results. First, I can get lower economic growth, similar to Liu et al. (2022), without requiring low interest rates which did not happen in Brazil (and other developing countries) on the same scale as in the US. I show both empirically and theoretically that the increase in skill concentration can happen due to higher high-skill labor supply. Second, I offer a potential mechanism to the observation made in Akcigit and Ates (2023) that there has been less knowledge diffusion in the US. Although ideas can be understood as public goods, turning ideas into productivity requires internal capabilities and skills (Cohen and Levinthal, 1989). By incorporating labor-dependent catching up in my model, I show how skill concentration lowers active R&D imitation by the laggard firm.

My paper also contributes to the recent literature on the rise of firm concentration. This rise, documented for developed countries, has been attributed to different reasons, including a decline in antitrust policy (Döttling, Gutierrez Gallardo and Philippon, 2017), technological change (Autor, Dorn, Katz, Patterson and Van Reenen, 2020, Olmstead-Rumsey, 2022, De Ridder, 2024) and diffusion (Akcigit and Ates, 2023), lower business dynamism (De Loecker, Eeckhout and Mongey, 2021), and demographics (Hopenhayn, Neira and Singhania, 2022). My contribution is in identifying a new channel through which high-skill concentration increases, i.e. increases in high-skill supply, as large firms benefit the most from an increase in human capital. This channel is particularly useful in the context of developing countries as several nations have experienced a large increase in high-skill supply (c.f. Figure A1 in the Appendix). My paper is more closely related to Olmstead-Rumsey (2022) and De Ridder (2024) in using an endogenous growth model to propose a novel mechanism that links the rise in market concentration to a decline in growth.[6] This paper, however, focuses on the role of human capital.

Finally, this paper makes an important contribution to the recent literature examining the effects of firm concentration in the labor market. Previous papers have linked firm labor market concentration with a reduction in wages (Dix-Carneiro and Kovak, 2015, Azar, Marinescu and Steinbaum, 2022, Benmelech, Bergman and Kim, 2022, Schubert, Stansbury and Taska, 2024, Felix, 2022), which we also see on the aggregate in Brazil specifically for high-skill workers. However, I provide causal estimates that the relationship between local skill concentration and the skill premium is non-monotonic, which I rationalize in a model of innovation. My contribution is, then, to show how conclusions can differ for high-skill labor when we take skill concentration into account. This paper is more closely related to Akcigit and Goldschlag (2023), which shows empirical evidence of inventor concentration at large firms in the US, and Manera (2022), which uses defensive R&D (vs. productive R&D) to explain inventor concentration at the sector level. Similar to my results, both studies show that leading firms face less incentives to implement new ideas once inventors are hired. However, while their results are based in leapfrogging models of innovation, I show how step-by-step firm interaction and labor market search are crucial to understanding the non-monotonic empirical trends with respect to economic growth.

The remainder of the paper is organized as follows. Section 2 describes the data. Section 3 explains the empirical strategy and shows estimation results along with robustness

[6]The high-skill labor channel is potentially related to Olmstead-Rumsey (2022) who finds a worsening of the quality of innovation done at small firms. If large firms are keeping the best inventors away from the labor market, or if small firms face hurdles in hiring high-skill labor for breakthrough innovation, then the resulting effect on growth is compounded.

checks. Section 4 rationalizes results with a step-by-step model of innovation with high-skill labor search. Section 5 uses the model to analyze counterfactuals and policy. Finally, Section 6 concludes.

# 2 Data

My main data source is the Annual Social Information Report (RAIS) which has annual, non-identifiable socio-economic data on employer-employee links from 1999 to 2017 in Brazil. RAIS contains data on location, type of employer, establishment size, establishment sector, job occupation, wages, work hours, duration of employment, and demographics including the worker's education level. All employers are required to send their employee data to the Ministry of Labor, which oversees RAIS, and face fines if they do not. As such, the database represents almost the entire labor force under formal employment, which is my main focus since I am interested in high-skill labor concentration. I do, however, exclude workers in the armed forces, police, firefighting, and politicians from the data, as these are not usually associated with a firm. Importantly, RAIS data identifies an establishment as an employer. This will be relevant when discussing the mechanism behind my results as establishments, in being a smaller constituent of a larger firm, face more intra-municipal competition. As most firms consist of a single establishment, I use "firm" and "establishment" interchangeably throughout the paper.

It is important at this point to specify a few definitions regarding workers and employers used in the empirical estimations. I define high-skill workers as those who have at least some undergraduate education, though they might not have finished their degree. This group corresponds to around 17.8% of workers in my sample. I use a broad definition of high-skill as it allows me to capture workers who have the capacity for productivity-enhancing activities regardless of their current occupation.[7] Particularly, it is not uncommon for a worker with a degree in an innovation-related field (e.g. engineering) to be hired in a non-innovative occupation (e.g. financial analyst). This worker, however, could still produce innovation if employed to do so. To show robustness of results, I also use a narrower definition of high-skill which I label "high critical-thinking workers." Specifically, starting with the set of aforementioned high-skill workers we narrow it down to those who are also employed in occupations requiring innovation-prone skills. Importantly, given the Brazilian context I consider throughout the paper the idea that

[7]This is further backed by Harrigan, Reshef and Toubal (2023) who shows how broadly defined "techies," i.e. engineers and technically trained workers, are important for innovation and technology adoption (vs. the narrow "scientists").

innovation includes not only frontier R&D but also imitation, technology adoption, and more incremental types of innovation. Data on skills by occupation comes from O*NET.[8] As for employers, I classify those with 500 employees or more as being large. In the case where a municipality does not have an establishment that matches this criterion, I consider as large those with 250 employees or more. If there are still no local large employers, I label as large those with 100 employees or more. Large employers, then, correspond to around 14.4% of all employer-employee links.[9]

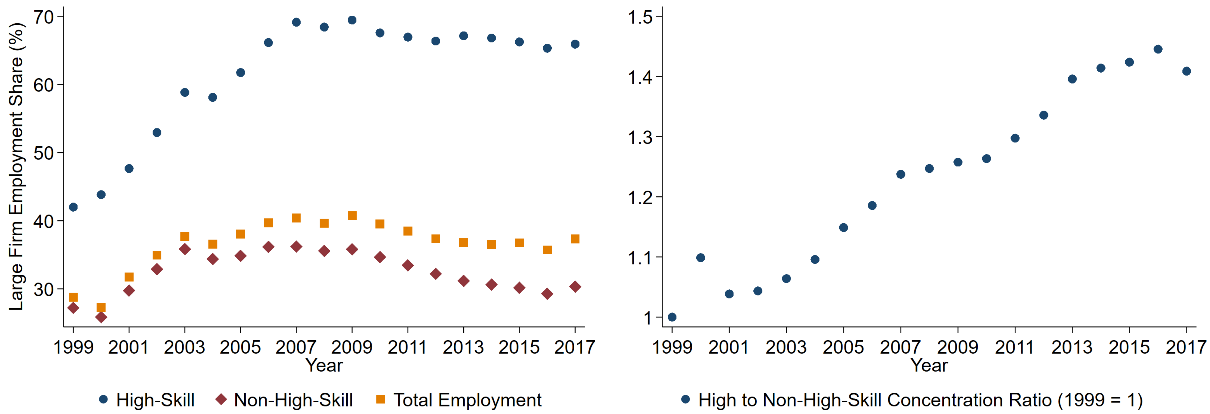
I also use the RAIS Estabelecimentos dataset which contains similar data to the employer-employee dataset collapsed at the employer level. This is useful as it allows me to calculate the municipal-level Herfindahl-Hirschman index (HHI) for total employment. This dataset, however, does not have variables separated by different types of employees such as low or high skill. As such, I am limited to calculating a firm-level, HHI-style measure of concentration for total employment, which is still useful as we can compare it with trends in high-skill concentration.

It is important to highlight that my results are specific to high-skill concentration. While my research design in Section 3 makes it clear that municipality-level results are specific to high-skill labor, we can also see in the aggregate data that high-skill concentration has followed a particular (rising) trend. I show this in Figure 1 which plots the evolution of high-skill, non-high-skill, and total employment concentration, calculated as the share of workers at large firms. We note one important fact: that only high-skill labor saw an important increase in concentration at large firms as non-high-skill concentration went up to then mostly decline. To quantify the relative increase in high-skill concentration, I show in Figure 1 the evolution of the ratio between high and non-high-skill concentration. This ratio increased by almost 50% relative to its value in 1999. This observation is robust to two other measures of skill concentration. First, we observe a similar rise in high-skill concentration with a HHI-based measure of concentration calculated using firm size bins. Second, there is also little change in the overall trend when we compare this bin-level HHI with a firm-level HHI, a comparison that we can only do for total employment due to data limitations. This provides evidence that the observed increase in high-skill concentration would remain the same had we been able to use firm-level high-skill employment shares. I show these different skill concentration measures in Figure A3 in the Appendix. Overall, these observations show that my results are not linked to a

[8]Formally, I define a high critical-thinking worker as those with at least some college education who are employed in occupations at the top skill quartile for one of the following: Math, Science, Critical Thinking, Active Learning, and Complex Problem Solving.

[9]Throughout the paper, I refer to firms not classified as large ones as small or non-large firms indistinguishably.

Figure 1: *Evolution of high-skill concentration and the high to non-high-skill concentration ratio*




Note: High-skill (non-high-skill) concentration is the median across municipalities of the local share of high-skill (non-high-skill) people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree.

rise in non-high-skill or total employment concentration, and that the mechanism that I propose in Section 4 is particular to high-skill labor.

Apart from RAIS data, I use data from the General Registry of Employed and Unemployed Workers (CAGED) to calculate the municipal-level net change in total number of workers. While there is no continuous panel data on municipal unemployment, the data from CAGED captures movements in the local unemployment rate. Similar to RAIS, the government requires firms to report the hiring and firing of formal workers, which is then imputed into CAGED. However, data points before 2020 do not systematically include temporary workers as firms were not required to report those. As I am mainly concerned with high-skill workers who are generally hired for full-time positions, this data aspect does not seem to be a problem. As such, for the municipality-year pairs where data is missing I input zeros which should be understood as no changes in the number of full-time formal workers.

I also use minimum wage data from IPEA to calculate nominal wages. This is necessary because RAIS provides wage data in units of the national minimum wage. I deflate nominal values using inflation data from the Brazilian Institute of Geography and Statistics (IBGE). Municipal-level population and GDP estimates are also obtained from IBGE along with data on the municipal share of informal workers which is available for 2000 through the census. Finally, state-level data on electricity consumption (in MWh), which I will use as a proxy for capital investment, comes from the Energy Research Office, a government-affiliated company. GDP and other firm-related variables are deflated using the GDP deflator.

For my difference-in-differences estimation, I use college creation and quality data from the National Institute of Educational Studies and Research (INEP). On the former, INEP has college-level data between 1999 and 2019. Although the data is non-identified for most years, we can identify any changes to the total number of colleges within a municipality. As for course quality, I rely on two national-level assessments called ENADE (National Student Performance Exam) and the CPC (Preliminary Course Score). ENADE is a test created in 2004 that most college students have to take to graduate which measures both general level knowledge and content that is specific to degree fields ("broad" and "specific," respectively).[10] Those taking the test must be at the end of their courses. The CPC is a composite indicator of quality, available since 2007, which takes into account the ENADE grade, teaching staff quality, student feedback, and an indicator of learning value added.

To construct my SSIV, I use loan-level data from the BNDES, which also includes information on borrower characteristics. This dataset covers the period between 2002 and 2017. It is worth it at this point to mention a few aspects of how the BNDES works. The development bank played an increasingly relevant role in the Brazilian economy throughout my sample, representing around 20% of total bank loans (or 10% of GDP) in 2015 according to the Brazilian Central Bank. The BNDES is mainly funded through taxes and the Treasury, and can offer loans to most firms that meet the criteria of its different loan products.[11] Most of its loans, designed to support national development and social causes, have below-market interest rates and facilitate investments in innovation, green technology, infrastructure, exports, among other areas. Loan eligibility criteria depend on firm size, sector, and the purpose of the loan. Its loan offer is heavily influenced by the Executive Branch of the national government, who can determine funding changes and pick the bank's CEO.

Finally, I use IBGE data from annual sector surveys for the sector-level balance tests with respect to my shift-share strategy. These are run for the following economic sectors: manufacturing, construction, retail and wholesale, and most services. Although data coverage varies in time and between surveys, I am able to compile supply-side data on revenues, value added, intermediate inputs, wages, and number of production workers.

I report summary statistics at the municipal-year level in Table A.1 in the Appendix.

[10] Though most colleges apply the test, it is only mandatory for private or federal universities.

[11] The BNDES is prohibited from offering loans to the banking, financial, weapons trade, adult entertainment, and gambling sectors.

# 3 Empirical Results

In this section, I first show reduced-form evidence using a difference-in-differences design that an increase in local high-skill supply in Brazil has led to a relative decline in GDP growth in municipalities where skill concentration was high. Using the same empirical design, I also show causal evidence that the increase in human capital led to higher high-skill concentration at large firms. I then present causal estimates for the relationship between local high-skill concentration and GDP growth using a SSIV. I also use the same SSIV to pin-down the effect of high-skill labor concentration on the skill premium. I then show that the shift-share design passes the recommended falsification tests in the literature and that results are robust to changes in the specification.

## 3.1 Difference-in-Differences Design: Increase in High-Skill Supply

We first look at how an increase in human capital led to lower GDP growth in Brazilian municipalities where skill concentration was high relative to places where it was low. Leveraging the same empirical strategy, I then show that the increase in high-skill supply led to an increase in skill concentration.

To identify the effect of high-skill supply on growth, I leverage data on college and university creation in Brazil between 1999 and 2019 from INEP. The college-education sector saw a boom since the late 1990's as a result of government policy. In particular, the 1996 reform which made it easier for institutions to set up courses and programs, the Higher Education Student Financing Fund (FIES) created in 1999 which offers subsidized loans to low-income students, and the 2004 College For All program (ProUni) which mainly offers college grants to low-income students from public schools. Figure A4 in the Appendix shows the strong increase in both the number of colleges and the share of college graduates in the population since the 1990's, particularly in the private sector. Importantly, by comparing the flat trends in the population share of graduates before the 1996 with the steady expansion in later years, it is clear that supply of colleges was being constrained by the legal framework in Brazil before the reform. Moreover, as the growing trend in the population share has yet to stop, we can infer that supply has yet to catch up with student demand. We can, then, exploit this substantial expansion of colleges to assess the effects of increasing high-skill supply on growth by comparing municipalities that received a new college in the period ("treated") to municipalities that did not ("control").

Before doing so, it is important to assess how these two groups differ. Both government entities and the private sector might well pick municipalities for new colleges based on specific characteristics that correlate with local economic growth. In particular, we can conceive that for-profit private colleges, which are around half of the private college group in the 2010's, choose municipalities where student demand is high. This could potentially threaten the identification assumption of the difference-in-differences which relies on the choice of municipality and timing of opening a new college to be as-good-as-random with respect to local growth. To assess this threat, I report in Table A.2 in the Appendix the summary statistics for both treated and control groups on different demographic and economic observables. Importantly, Table A.2 shows that treated and control groups do differ on observables. In particular, places picked for new colleges are wealthier and more populous. On the other hand, groups look similar with respect to share of workers employed in services and educational profile. Though relevant, these differences between the two groups of municipalities do not constitute an impediment *per se* to using untreated localities as a control group so long as the choice of where to place a new college was not correlated with local growth trends.

For the case of Brazil after the 1996 reform, the choice of municipality for a new college is *ex-ante* likely to be as-good-as-random with respect to local growth. This is because the supply of colleges was suddenly unleashed in 1996. As such, new colleges from that point onward faced significant excess demand, increasingly so after the government launched the subsidy programs in 1999 and 2004. This implies that demand was not the differential factor in choosing one municipality rather than another. Instead, marginal factors such as having the support of a local politician or having ties with the local economy became the determining elements behind the choice of where to build a new college.[12] Finally, it is important to highlight that setting up a new college takes years from idea to inauguration, mostly because any new college needs to be approved by the federal government, a step that can take up to three years. This makes targeting particular trends in local observables an unreliable procedure.

We can check the data for evidence that our assumption of municipality choice being as-good-as-random holds. I do so in several ways. First, I check for pre-trends to see whether new colleges target high- or low-growth places, both on growth and employment flows. Second, I assess pre-trends on proxies for local demand and competition to check whether the data supports our idea that local demand did not play a significant role in choosing a municipality because it was high everywhere. Third, I construct a placebo group of municipalities by matching treated and control groups on population

[12] Anecdotal evidence includes a local mayor donating land, a representative using their budget allocation to support the construction of a new campus, and educators with local ties taking advantage of the 1996 reform to open a college.

level, share earning minimum wage or lower, share who only completed the $5^{th}$ grade, unemployment rate, and illiteracy rate, all in 2000. I show in Table A.2 the summary statistics of both the placebo group and the matched treated municipalities. Although both groups still differ on population size, all other observables are a close match which allows me to test whether differences in observables can explain my results. Fourth, instead of comparing places that received a new college with places that did not, I can compare the former with places that received a college in the last year of my sample. With the exception of the year when they are treated, last-treated municipalities provide a valid comparison group with the treated subsample (Sun and Abraham, 2021). Finally, I can also compare treated municipalities with those that, at a given time, have not yet received a new college, though will get one in future years. This not-yet-treated group also provides a valid comparison group (Callaway and Sant'Anna, 2021).

Having defined the empirical strategy, I identify the effect of an increase in high-skill supply on local growth by running the following difference-in-differences specification for municipality $i$ at time $t$:

$$Y_{i,t} = \sum_{\substack{k=-7 \\ k\neq -1}}^{17} \mathbb{1}_{\{D_{i,t}=k\}} \left[ \beta_{1,k} \mathbb{1}_{\{HSConc_{i,init} \leq p\}} + \beta_{2,k} \mathbb{1}_{\{HSConc_{i,init} > p\}} \right] + \alpha_i + \delta_t + \epsilon_{i,t} \tag{1}$$

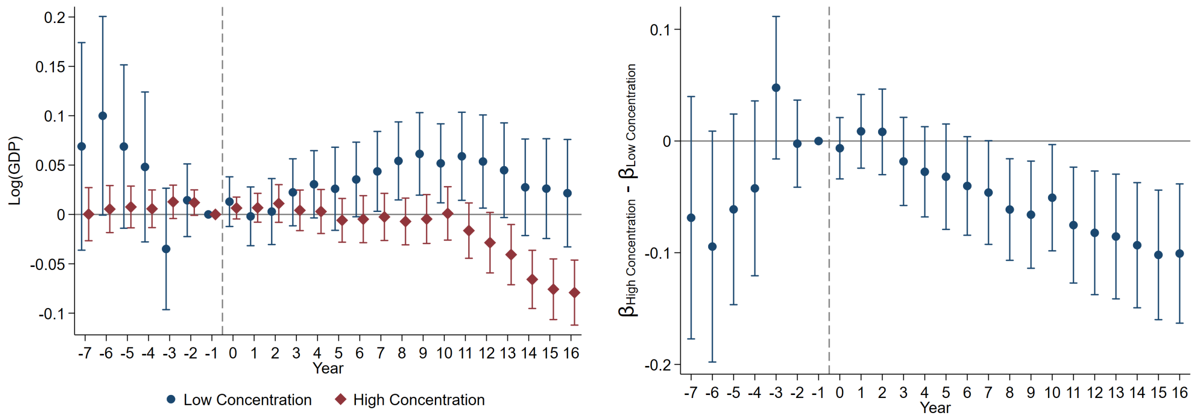
where $Y_{i,t}$ is the logarithm of GDP per-capita, $D_{i,t}$ is a binary treatment which is equal to one if a college or university were created at municipality $i$ at time $t$, $HSConc_{i,init}$ is the initial high-skill concentration,[13] $p$ is is the percentile threshold that defines both low and high concentration municipalities, and $\delta_t$ and $\alpha_i$ are time and municipality fixed-effects, respectively.[14] Skill concentration is defined as the sum of high-skill workers in large firms divided by the total number of local high-skill workers. Importantly, Equation 1 allows us to capture whether high-skill concentration plays a role in the effect of human capital supply on local growth by comparing $\beta_{1,k}$ and $\beta_{2,k}$. Notice also that in using the logarithm of local GDP, estimates can be interpreted, approximately, as the difference in long-term growth between treated and control municipalities.

We can identify the set of $\beta_{1,k}$ and $\beta_{2,k}$ from the assumption of parallel trends. As aforementioned, the intuition behind identification is that the decision and timing of creating a new college are unrelated to local growth trends. We can assess evidence supporting

[13]For untreated municipalities, the initial skill concentration is the average concentration in the first two periods for which I have data. For treated units, I use the average concentration in the three years prior to treatment. Results are robust to varying this time window.

[14]While some municipalities report receiving new colleges multiple times, I consider treatment timing to be the year when a municipality reports receiving its first new college.

Figure 2: *Difference-in-differences estimates of the effect of college creation on local growth at municipalities with high and low high-skill concentration*




Note: High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Low (high) concentration municipalities are defined as those below (above) the $14^{th}$ percentile of high-skill concentration averaged within the three-year period pre-treatment, for those treated, or below (above) the same concentration level averaged in the initial two periods for the non-treated so as to get similar threshold levels in both cases. Vertical bars represent the 90% confidence interval.

this assumption by looking at pre-trends between treated municipalities and the control group. I estimate Equation 1 and plot the set of $\beta_{1,k}$ and $\beta_{2,k}$ in Figure 2 along with the difference $\beta_{2,k} - \beta_{1,k}$ between high and low skill concentration municipalities. We notice three important points. First, results show evidence of no pre-trends for both groups of treated municipalities.[15] Second, point-estimates are close to zero in the first three years of treatment, while the first significant coefficient only happens a few years latter for low concentration municipalities. Third, the difference between low and high skill concentration intensifies in time. All three provide initial support for our identification strategy. The first point is reassuring as it is what we would expect if the choice of where to open a new college is unrelated to local growth trends. The second point is in line with the fact that it takes around four years for the first student cohort to graduate, hence we should not expect a significant effect on growth in the early years after treatment. Finally, we should expect results to evolve in time as further cohorts add to the local supply of human capital, as shown in Figure A5 in the Appendix.

Results show heterogeneity in the effect of human capital on local growth. After the creation of a new college, we observe a positive and significant effect on local growth in municipalities where high-skill concentration was low before treatment. This effect, however, is mute at municipalities where skill concentration was high and, surprisingly,

[15]P-values for the joint test of significance: 0.11 (low concentration) and 0.72 (high concentration).

negative in the long term as growth declines. This highlights how an increase in human capital can result in a decline in growth depending on the local level of skill concentration. The difference between both sets of coefficients is significant (and negative) at 10% significance level from eight years after treatment onwards. The estimated relative decline in long-term growth at places with elevated high-skill concentration is around 10%, or a 0.9% average relative decline in yearly growth from the incremental increase in local skill supply, which effectively starts in $t + 4$ after the first cohort graduates. While the long-term effect from a new college on growth looks remarkable, it is important to notice that in most places the increase in the local supply of skills is relatively quite significant. As I show in Figure A5, after around 10 years since the first cohort graduates the local supply of high-skill workers increases, on average, by around 90% relative to the pre-treatment average. Finally, I show in Figure A6 in the Appendix that the significant decline in long-term growth is also captured in a specification without the heterogeneity by the level of skill concentration as, across all municipalities, we observe a decline in growth of around 6%. This is evidence that results are not being driven by confounding variables related to the level of skill concentration.[16]

It is important at this point to assess the evidence for the identification assumption. As aforementioned, my estimation relies on the choice of municipality for a new college being as-good-as-random with respect to local growth.[17] While it is reassuring to find no significant pre-trends in Figure 2, we can check the data for further evidence. First, I show in Figure A7 that we also do not observe significant pre-trends in the local stock of formal employees.[18] Second, I show in Figure A8 that we also fail to reject the parallel trends hypothesis on both the population share of the graduating college cohort and the difference between the number of new high-school graduates and the incoming first-year college cohort. Both represent different ways to gauge local demand for college education. The lack of significant pre-trends corroborates our intuition that local demand was not a major factor in determining where to open a new college as supply was severely constrained prior to 1996 and has yet to catch up by the end of my sample.

We can also check the robustness of our results, and the identification assumption, to

[16] Municipalities where skill concentration is high or low look similar on observables, as I show in Table A.3.

[17] It is important to highlight that for the results on the difference between high and low skill concentration coefficients, confounders would need to correlate not only with local growth but also with local skill concentration to be able to affect results, to the extent that if high and low skill concentration coefficients are equally biased, the difference will cancel out the bias. As such, the identifications strategy for the difference in coefficients is stronger.

[18] While we do not observe negative and significant estimates on employment, this is likely due to shifts between formal and informal sectors.

changes in the sample. The intuition behind this exercise is that if results remain robust in settings where threats to identification are lower, then we have evidence that such threats are not driving our results. I start by showing in Figure A9 the estimation results using the matched subsample where treated and control observations are matched on observables. Results on the difference $\beta_{2,k} - \beta_{1,k}$ are similar to the baseline, which is evidence that the differences in covariates reported in Table A.2 are uncorrelated with local GDP growth. In absolute levels, however, results using the matched sample are overall higher than the ones in Figure 2 as we now do not observe a significant decline in growth at municipalities with higher skill concentration. Nonetheless, this difference in results can be explained with the model I introduce in Section 4. The reason for a lack of decline in growth is that the restricted, matched sample has a lower average high-skill concentration level. Hence, the increase in skill concentration from a more human capital supply is not large enough, in this sample, to induce a decline in growth.[19]

Finally, we can assess our identification strategy by using different control groups. As with the previous exercise, if results are robust in a setting where the identification assumption is slightly different, we have evidence that our assumption is valid. I do this in two ways by switching the untreated control group with either the last-treated cohort or the not-yet-treated observations, both of which have been shown to provide valid comparison groups (Sun and Abraham, 2021, Callaway and Sant'Anna, 2021, de Chaisemartin and D'Haultfœuille, 2024). In both cases, treated and control municipalities will receive a new college at some point in my sample. Hence, the identification assumption is now on the timing of receiving a new college such that treatment assignment between early and later municipalities looks as-good-as-random.[20] This is likely since municipalities look similar on observables as shown in the summary statistics in Tables A.4 and A.5 for last-treated and not-yet-treated groups, respectively. Anecdotal evidence also corroborates the assumption on the timing as new colleges take, on average, many years to be created as founders need government approval, appropriate facilities and staff, and a procedure to formally enroll students. All these steps can take different amounts of time for reasons that are unrelated to local growth.

Results using last-treated and not-yet-treated are similar to baseline estimates. Starting with the former, I show in Figures A10 and A11 results for the individual coefficients and the difference in effect between high and low skill concentration municipalities. Although noisier, last-treated estimates are in line with baseline ones. I then show in Figure A12 results using the not-yet-treated group as control. As the estimation is noisier and

[19]C.f. Section 5 for a detailed explanation of the mechanism behind these results.

[20]Similar to the assumption made in Nimier-David (2023).

requires calculating estimates between different municipality cohorts, I make two important modifications. First, I set the threshold between low and high to be the $20^{th}$ percentile to reduce standard errors. Second, I set the treatment period $t = 0$ to be the period when the first student cohort is expected to graduate. This increases the number of not-yet-treated observations, though both changes render estimate by estimate comparisons with the baseline estimation difficult. Nonetheless, results look qualitatively similar, and we can visually infer that the difference between coefficients in high and low skill concentration places is significant and negative. Hence, both last-treated and not-yet-treated estimations show evidence in support of the validity of the identification assumption of the baseline estimation.

I show further robustness of results to changes in the sample or in the specification. On the former, estimates are robust to restricting the sample to municipalities that do not have colleges, either in all periods (control) or in the pre-treatment period (treated), as shown in Figure A13. This is reassuring as we might worry that including municipalities with pre-existing colleges could bias results if these colleges expand their student intake in response to college creation elsewhere. As results stay the same, this effect does not affect baseline estimates significantly. On the later, I first show that results remain the same if we increase or decrease the threshold $p$ that defines a high or low high-skill concentration municipality, as shown in Figures A14 and A15 for $p = 12\%$ and $p = 17\%$, respectively. I then show in Figure A16 that results remain unchanged if we add to the specification dummies for the leads and lags of municipalities that reported receiving new colleges twice or three times. Finally, results are robust to running a weighted specification where we weight by the log of local population, as I show in Figure A17. This is evidence that results are not being driven by the direct economic effect of new colleges on local GDP, mainly because new colleges represent little of the local economic activity. Importantly, in all cases we find evidence of no significant pre-trends. This confirms the *ex-ante* intuition that targeting high or low growth municipalities when opening a new college is unlikely as several other factors come into play.

Finally, results are unchanged if we use estimators robust to heterogeneous treatment effects and non-binary treatment. The literature on difference-in-differences estimators has shown that estimates can be biased in the presence of heterogeneity in treatment effects (c.f. Roth, Sant'Anna, Bilinski and Poe, 2023 for a summary). Moreover, in our context it is possible that more than one college is created within a single municipality over time, which ultimately constitutes multiple treatments. We can, then, assess whether alternative estimators that take into account such cases give different estimates. For the case of heterogeneous effects, I show in Figure A18 results using the estimator proposed in Sun and Abraham (2021) that restrict the control group to never-treated units, avoiding the issue of "forbidden comparisons" which may bias estimates. Assuringly, results remain indistinguishable from baseline ones. As for the possibility of non-binary, I show in Figure A19 estimates using the estimator proposed in de Chaisemartin and D'Haultfœuille (2024), which aggregates the treatment effect of municipalities experiencing different treatment paths. Once again, estimates are quite similar to baseline ones.

Hence, the relationship between human capital and local economic growth depends crucially on skill concentration. Similar increases in skill supply affect municipalities differently whether they have high or low high-skill concentration at large firms. In particular, highly concentrated places perform worse, showing a long-term decline in growth. Having established this empirical result, we now investigate the underlying mechanism that explains this heterogeneity in growth by shifting our focus to local high-skill concentration.

We start by noting the significant rise in local skill concentration in Brazil since 1999. As shown in Figure A2, high-skill concentration increased around 25 percentage-points between 1999 and early 2010s. Naturally, this trend could have different causes. There is a rich literature, mostly on developed countries, linking different mechanisms to a rise in firm total concentration, either measured in terms of revenue or total employment.[21] While some of the proposed explanations may apply to Brazil in the same period as my analysis and might explain a rise in high-skill concentration, I propose adding a new driver which is the increase in high-skill supply. The intuition behind this channel, which I formalize in Section 4, is that large firms are the ones who mostly benefit from the additional supply of high skill, increasing their gap relative to small firms.

To identify the effect of high-skill supply on concentration I leverage the same empirical strategy as the one I used to pin-down the effect of skill supply on GDP growth. That is, I run a specification that is similar to Equation 1. For municipality $i$ at time $t$:

$$\Delta HSConc_{i,t} = \sum_{\substack{k=-7 \\ k\neq -1}}^{17} \beta_k \mathbb{1}_{\{D_{i,t}=k\}} + \alpha_i + \delta_t + \nu_{i,t} \tag{2}$$

where $\Delta HSConc_{i,t}$ is the cumulative growth rate of local high-skill concentration between the first period for which we have data and $t$.

Similar to Equation 1, the identification of the parameters of interest $\beta_k$ requires an assumption on municipality choice for a new college. Consistent estimation of Equation 2 relies on the assumption that both control and treated municipalities would have experi-

[21] C.f. Section 1.

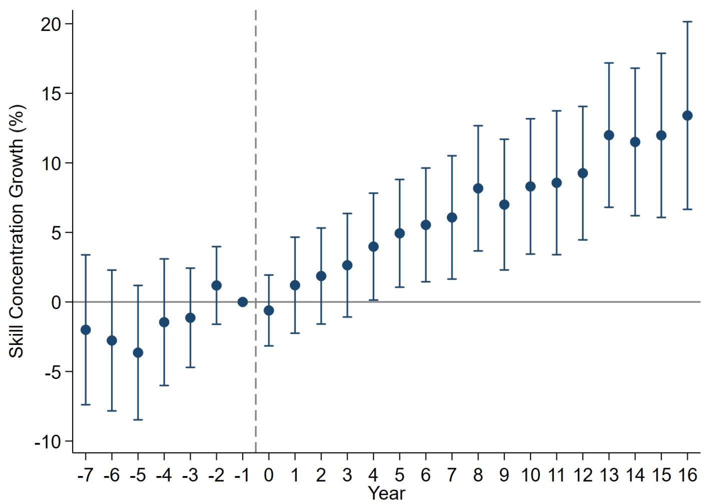
enced similar trends in local skill concentration across firms had there been no treatment. As mentioned before, these two groups do look different in some parameters (Table A.2). However, we can apply a similar reasoning to the used for GDP growth: any imbalance between groups is not a problem so long as the choice of where to build a new college is unrelated to local trends in high-skill concentration at large firms. While I provide evidence of the validity of this assumption, it makes *a priori* sense that it holds for the same reasons discussed in the specification using local GDP. Along with checking for the presence of pre-trends, I also assess the presence of imbalance by running a robustness check using the matched sample of placebo and treated municipalities.

We can then proceed with estimating the set of $\beta_k$. I show results in Figure 3. As with our results on growth, Figure 3 highlights a few reassuring points. First, there is no evidence of pre-trends, which is in line with the assumption that both treated and control groups would have behaved similarly in the absence of treatment. Second, we observe a similar delay, relative to treatment period, in significant results as the time between starting college and graduation takes on average four years. After this initial period, however, results are significant and show an increase in high-skill concentration in large firms due to college creation and the increase in the flow of high-skill people. The magnitude of the increase is also important as it represents a rise of around 12% in concentration a decade after the first students start graduating.

Results on high-skill concentration are robust to different specifications, changes to the sample, and alternative estimators. First, I show that estimates remain similar if we use the matched placebo group as our control group, as shown in Figure A20 in the Appendix, evidence that any imbalance between treated and control groups is not affecting estimates. Second, results are also robust to restricting the sample to municipalities that do not have colleges, at all (control) or prior to treatment, as shown in Figure A21. Third, I show in Figure A22 that results are robust to adding controls for the leads and lags of municipalities that reported receiving new colleges twice or three times. Fourth, Figure A23 shows that results are robust to using a HHI-based measure of local high-skill concentration. Finally, results remain unchanged if we use instead the robust estimator proposed in Sun and Abraham (2021), as shown in Figure A24, or the estimator proposed in de Chaisemartin and D'Haultfœuille (2024), as shown in Figure A25.

Evidence, then, points towards an important role of the steep increase in high-skill supply in Brazil in the rise in local skill concentration. To properly identify the effect of high-skill supply, we had to focus on college creation which limits how we can translate

Figure 3: *Difference-in-differences estimates of the effect of college creation on local high-skill concentration*




Note: High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Sample excludes observations with no workers, high-skill or not, at large firms. Vertical bars represent the 95% confidence interval.

those results to the aggregate economy due to the missing intercept problem.[22] Nonetheless, we can proceed with a back-of-the-envelope calculation to gauge the magnitude of this high-skill supply channel on high-skill concentration. If our conclusions on college creation can be applied broadly to the rise of college graduates, whose numbers more than tripled between 2000 and 2010, the 12% increase in high-skill concentration could potentially explain almost half of the average national increase in concentration between 2000 and 2010.[23] Although a simplification, this calculation shows that the high-skill supply channel seems quite relevant in explaining the increase in local skill concentration.

### 3.2 Shift-Share Design: From Skill Concentration to Growth

After showing the causal link between human capital supply and high-skill concentration, we now proceed with the second step of the skill concentration channel. That is, the relationship between local skill concentration and economic growth.

[22]For instance, factor mobility can complicate translating the difference-in-differences estimates to the aggregate economy.

[23]To arrive at this conclusion, we start by noting that the 12% increase in local skill concentration is associated with a rise of around 2 percentage-points in the local share of high-skill people (Figure A5). We can then extend our result by assuming that the 4.5 percentage-point increase in the national share of high-skill people in the same period (Figure A2) caused a similar growth in high-skill concentration as the one measured for new colleges. Finally, we compare this number to the around 24 percentage-point increase in the aggregate concentration.

Results from Section 3.1 suggest the relationship between local high-skill concentration and local growth is non-monotonic. The reason for this is the following. We have, so far, seen that the increase in the supply of human capital has caused both an increase in local skill concentration and has had a heterogeneous effect on local growth depending on the level of this concentration. If high-skill concentration plays a role in connecting human capital and growth, then its increase should lead to different effects on local growth depending on whether it the level of skill concentration is high or low. In particular, given an increase in skill concentration we expect local growth to rise in low concentration municipalities and to fall in high concentration ones.

We, then, proceed to test this hypothesis by looking at how high-skill concentration at large firms affects GDP growth rates at the municipal level. Our goal is to assess whether the relationship between these variables is non-monotonic. As I do not want to impose a functional form *a priori*, for municipality $i$ at time $t$ the main specification is the following:

$$
\begin{aligned}
y_{it} = {} & \beta_1 HSConc_{i,t-1} \mathbb{1}\{HSConc_{i,t-1} > p\} + \beta_2 HSConc_{i,t-1} \mathbb{1}\{HSConc_{i,t-1} \leq p\} + \\
& \gamma X_{i,t-2} + \epsilon_{it}
\end{aligned}
\tag{3}
$$

where $y_{it}$ is real GDP per-capita growth, $HSConc_{i,t-1}$ is the concentration of high-skilled labor at large firms, $p$ is a specific percentile threshold, and $X_{i,t-2}$ are controls which include time and municipality fixed-effects, and a constant for $\mathbb{1}\{HSConc_{i,t-1} > p\}$.[24] As before, high-skill concentration is defined as the sum of high-skill workers in large firms divided by the total number of local high-skill workers. I use lagged high-skill concentration as my main regressor to account for the delay between the hiring decision and actual employee deployment, including any new-hire training. Effectively, Equation 3 estimates two slopes: one for municipalities where high-skill concentration is relatively low ($\beta_1$) and one for places where it is relatively high ($\beta_2$). We can then compare the signs of $\beta_1$ and $\beta_2$ for evidence of non-monotonicity. If our hypothesis holds, we expect $\beta_1$ to be positive while $\beta_2$ is negative.

As high-skill concentration in a municipality can depend on other endogenous variables and be affected by local GDP growth rates, I address endogeneity concerns via an instrumental variable approach. A possible issue with estimating Equation 3 is that a municipality experiencing high growth could be seen by entrepreneurs as a good place to start (or expand) a company. This, in turn, may affect high-skill concentration at large incumbents, biasing my results. Municipality-specific confounders such as local productivity changes can also pose a threat to identification. I, then, propose a shift-share IV to

[24] Controls are twice lagged to match the timing of the shocks in the SSIV.

address this endogeneity. The SSIV is constructed by leveraging heterogeneous exposure to public loans from the BNDES. As explained in Section 2, the BNDES loan portfolio, both in terms of size and client characteristics, is heavily influenced at the national level. While local demand for public loans is affected by local supply and demand conditions, I assume exogeneity relative to Equation 3 of changes to the sector-level, national loan amount offered by the BNDES in any given year. This identification strategy consists of the "shift-approach" discussed in Borusyak, Hull and Jaravel (2021). As such, I can use the heterogeneity in local-level exposure, measured by sector employment shares in both large and small firms, to national changes in loan offer to estimate Equation 3 consistently.

Specifically, I instrument Equation 3 with the following SSIV:

$$B_{i,t-2} = \sum_{n} s_{in,t-3} g_{n,t-2} \tag{4}$$

where $g_{n,t-2}$ is the sector $n$ shock ("shift") at time $t-2$, defined as the growth rate of the national loan amount, and $s_{in,t-3}$ is the exposure of each municipality $i$ to sector $n$'s shock at time $t-3$, measured as the local-level high-skill employment share in sector $n$. The SSIV is one-period lagged relative to the endogenous variable to account for the timing between loan issuance and actual spending, and I use the 2-digit Brazilian National Classification of Economic Activities (CNAE) to classify the $n = 1, ..., N$ sectors.

Following Borusyak et al. (2021), the validity condition for the shifts can be written as:

$$E\Big[\sum_{t}\sum_{i} B_{i,t-2}\epsilon_{it}\Big] = E\Big[\sum_{t}\sum_{i} \epsilon_{it} \sum_{n} s_{in,t-3} g_{n,t-2}\Big] = E\Big[\sum_{t}\sum_{n} \bar{\epsilon}_{n,t} s_{n,t-3} g_{n,t-2}\Big] = 0 \tag{5}$$

where $s_{n,t-3} = \sum_i s_{in,t-3}$ and $\bar{\epsilon}_{n,t} = \frac{\sum_i s_{in,t-3}\epsilon_{it}}{\sum_i s_{in,t-3}}$. Equation 5 shows how we can rewrite the orthogonality condition with respect to the SSIV $B_{i,t-2}$ as a condition on the orthogonality of shocks $g_{n,t-2}$. Intuitively, the validity condition assumes national shocks are uncorrelated with municipality-level confounders and do not systematically favor certain industries in a way that may bias results. We assess this point through falsification tests in Section 3.4.

Before proceeding with the estimation, it is important to split the SSIV between large and small firms. Using $B_{i,t-2}$, calculated by bundling all firm sizes together to instrument for high-skill concentration, is problematic as loans to large and small firms affect local large firm concentration differently. It is reasonable to expect a national increase of loans to large firms to increase high-skill concentration in these firms as they increase in size from the BNDES boost, while an increase in small firms loans may have the opposite effect. As such, it is important to separate shocks to national loan levels to both large and non-large firms. Moreover, given that the BNDES has different loan programs and conditions for large and small firms, there is enough variation by firm size to warrant a split up of the instrument between large and non-large firms. Effectively, I separate shocks $g_{n,t-2}$ and shares $s_{in,t-3}$ of small and large firms as if they were from different sectors, and calculate two shift-share instruments: $B_{i,t-2,large}$ and $B_{i,t-2,small}$.[25] I then instrument $HSConc_{i,t-1}$ in Equation 3 with both SSIVs, each interacted with $\mathbb{1}\{HSConc_{i,t-1} > p\}$, totaling four instrumental variables.

One concern in using these instruments is that they may affect other firm-level inputs which could confound the effect of high-skill concentration. For instance, an increase in national BNDES funds for large firms increases not only high-skill hiring but also non-high-skill labor and investments in capital, both of which would increase revenues and affect GDP growth. While capital takes longer to adjust than high-skill labor, changes in non-high-skill labor are a potential issue. To deal with the latter, I add total non-high-skill hiring as one of the controls in Equation 3. To address the same endogeneity issue as with high-skill labor concentration, I instrument non-high-skill hiring with the same set of instruments used for high skill concentration.[26] To further highlight the particular role of high-skill concentration relative to concentration based on total number of workers, I also add one specification where I control for the municipal-level employment HHI measure of concentration, similarly instrumented with the available IVs.

Even though capital formation takes longer than labor hiring, we can still worry about a violation to instrument validity from capital investing. As there is no data available on total capital stock at the municipality level, I proxy investment with changes in electricity consumption.[27] Although electricity consumption data is only available at the state-level, I get municipality-level variation by multiplying the per-firm, per-worker consumption with the local number of firms.[28] I then divide local consumption by local GDP and add the change in local electricity consumption as a control variable, which I treat as

[25] Shocks are winsorized at the $3^{th}$ and $97^{th}$ percentiles to avoid results being driven by abnormal relative increases in shocks.

[26] We may worry that controlling for non-high-skill hiring might introduce bias in the estimation via a "bad control" problem. I report Monte Carlo simulations in Section A.3 in the Appendix that validate the identification strategy.

[27] An idea that goes back to Taylor (1967), Moody (1974), and Burnside, Eichenbaum and Rebelo (1995).

[28] I use state-level electricity consumption of the manufacturing sector as it likely correlates more strongly with capital services. Results are unchanged if, instead, I use total consumption of both manufacturing and services.

endogenous and instrument with the IV set, in Equation 3.[29]

I report 2SLS results for Equation 3 in Table 1. I set the threshold $p$ between high and low concentration to the $16^{th}$ percentile to maximize IV relevance.[30] Column (1) only includes time and municipality fixed-effects while Column (2) adds local-level controls. In Column (3) I add the 2000 local informality share interacted with year fixed-effects as a control. Columns (4)-(8) assess the potential bias threat from non-high-skill hiring (4), new capital formation (5), total employment concentration (6), and both non-high-skill hiring and capital formation (7-8). I add a third SSIV in Columns (5)-(8) which is the same as the one described in Equation 4 except that I do not separate shocks to large and small firms and I use total employment share (vs. high-skill employment shares) as my SSIV exposure shares. As joint instrument relevance declines when we add all endogenous variables, I run in Column (8) the same specification as in Column (7) using the Limited Information Maximum Likelihood (LIML) estimator instead of 2SLS as the LIML estimator has lower small sample bias due to weak instruments.[31] Joint F-statistics are above the usual weak-IV threshold in all specifications except (7) and (8) though the negligible change in estimates between 2SLS and LIML suggest a small bias. I assess this point further by reporting the Olea-Pflueger effective F-statistics (Olea and Pflueger, 2013) for $HSConc_{i,t-1}\mathbb{1}\{HSConc_{i,t-1} > p\}$ and $HSConc_{i,t-1}\mathbb{1}\{HSConc_{i,t-1} \leq p\}$ separately, along with the respective critical values at significance level 5% and a 10% "worst-case" bias. F-statistic values are above the critical values in all specifications. Finally, we do not reject the null for the J-test of over-identification. This provides initial support for the validity of the instruments, a point which I analyze further for the SSIV in Section 3.4.

Results in Table 1 show a non-monotonic relationship between high-skill concentration and local GDP growth that corroborate our initial hypothesis. In all specifications, an increase in high-skill labor concentration at large firms increases local GDP growth in places where this concentration is low to begin with (positive coefficient). This effect, however,

[29]Table A.6 in the Appendix shows that the SSIV leads, as expected, to more hiring of high-skill and high critical thinking workers. The effect on non-high-skill hiring is different whether loans are for large or small firms. Results also indicate that the SSIV using both small and large firm loans lowers the ratio of per-worker electricity consumption over GDP, evidence that capital formation is not happening at significant levels.

[30]Though relevant for the estimation, the particular choice of threshold $p$ does not matter for the conclusion on non-monotonicity. Ideally, if we call $f(X)$ the true function relating dependent and independent variables, we want to pick a value of $p$ that is close to the point where $f'(X) = 0$, i.e. a local minimum/-maximum. I show robustness to the choice of $p$ in Section A.4.

[31]While LIML is known to be inconsistent under heteroskedasticity and many instruments, the bias is small when the number of instruments is $< 10$ (Hausman, Newey, Woutersen, Chao and Swanson, 2012). Importantly, all specifications include the local sum of shares $s_{i,t-2} = \sum_n s_{in,t-2}$ as recommended in Borusyak et al. (2021). Results using the heteroskedastic version of LIML are unchanged.

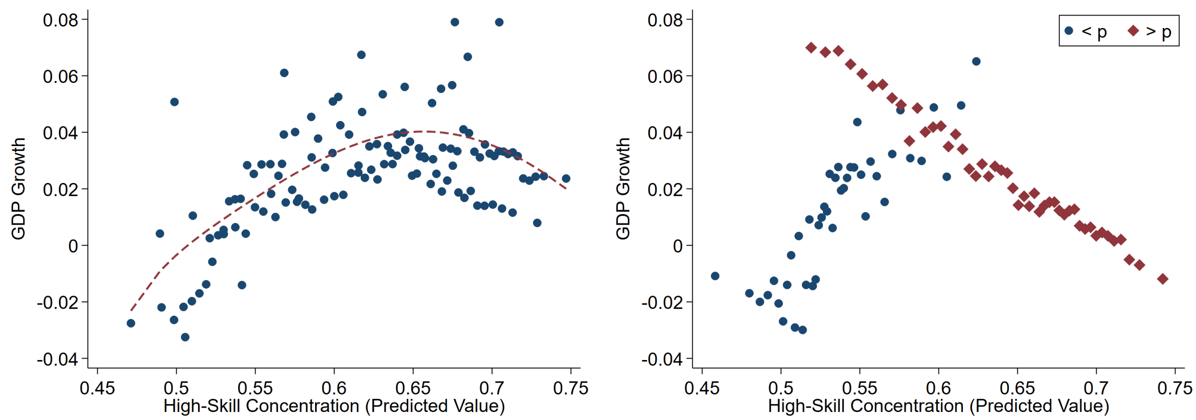
Table 1: Effect of high-skill concentration in large firms on local GDP growth in places with high and low concentration

| | GDP Growth | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=0 \times HS\ Conc._{t-1}$ | 0.973*** | 0.903** | 0.972** | 1.275*** | 0.889** | 0.876** | 1.291*** | 1.288*** |
| | (0.286) | (0.287) | (0.301) | (0.360) | (0.284) | (0.292) | (0.380) | (0.378) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=1 \times HS\ Conc._{t-1}$ | -0.489*** | -0.489*** | -0.499*** | -0.570*** | -0.489*** | -0.481*** | -0.573*** | -0.572*** |
| | (0.109) | (0.108) | (0.112) | (0.123) | (0.108) | (0.110) | (0.125) | (0.125) |
| N | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Informality | | | Yes | | | | | |
| Non-High-Skill | | | | Yes | | | Yes | Yes |
| Capital Proxy | | | | | Yes | | Yes | Yes |
| Employment HHI | | | | | | Yes | | |
| LIML | | | | | | | | Yes |
| Joint F-statistic | 32.4 | 32.6 | 28.9 | 13.0 | 14.6 | 25.7 | 6.0 | 6.0 |
| J-test, p-value | 0.12 | 0.13 | 0.17 | 0.94 | 0.18 | 0.15 | 0.99 | 0.99 |
| OP F-statistic, $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=0$ | 38.0 | 38.2 | 35.2 | 33.8 | 37.2 | 36.1 | 30.7 | 30.7 |
| OP Critical Value, $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=0$ | 16.7 | 16.8 | 16.5 | 15.5 | 16.5 | 16.5 | 14.8 | 18.2 |
| OP F-statistic, $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=1$ | 41.8 | 41.4 | 37.7 | 48.9 | 39.5 | 38.9 | 47.3 | 47.3 |
| OP Critical Value, $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=1$ | 13.4 | 13.4 | 12.7 | 5.2 | 12.1 | 12.3 | 4.9 | 12.9 |

High-skill concentration ($HS\ Conc$) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. High-skill concentration threshold $p$ is set at the $16^{th}$ percentile. GDP Growth is real per-capita local GDP growth winsorized at the $1^{st}$ and $99^{th}$ percentiles. Columns (1)-(4) use $B_{i,t-2,large}$ and $B_{i,t-2,small}$ as instruments, each interacted with $\mathbb{1}\{HSConc_{i,t-1} > p\}$. Columns (5)-(8) add to the IV set a SSIV calculated using both small and large shocks together, and using total employment shares as exposure shares. All specifications control for the lagged local sum of shares $s_{i,t-2}$. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. Informality refers to the 2000 ratio of informal workers over total employment interacted with year fixed-effects. Non-High-Skill refers to the total non-high-skill hiring instrumented with the SSIVs. Capital Proxy refers to the capital proxy variable calculated using the change in local electricity consumption instrumented with the SSIVs. Employment HHI refers to the HHI measure of concentration calculated for total employment. J-test refers to the overidentification test in Hansen (1982). OP F-statistic and Critical Value refer, respectively, to the Olea-Pflueger effective F-statistic and the critical value for a 5% significance level and a 10% "worst-case" bias. Municipality-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

reverses (i.e. a negative coefficient) at places where labor was already highly concentrated at the large firms. This is in line with our initial hypothesis that local skill concentration is the key channel between human capital supply and local growth given our results in Section 3.1. Results are significant in all specifications and show that a growth slowdown can be induced by an accumulation of high-skill labor at large firms. Controlling for either (or both) non-high-skill hiring and our proxy for capital formation does not change coefficients significantly which indicates that potential biases from changes in other inputs

Figure 4: *Binned scatter plot between local GDP per-capita growth and the first-stage predicted values, both unconditional (left) and conditional on being below or above the threshold p (right)*




Note: Plots show the predicted value of the $1^{st}$ stage of the 2SLS ($\widehat{HSConc}_{i,t-1}$) on the x-axis. As the threshold $p$ is defined over high-skill concentration ($HSConc_{i,t-1}$), right-hand side figure shows separate binned scatter plots for observations below or above the threshold $p$. Plots were made using the procedure in Cattaneo, Crump, Farrell and Feng (2024), controlling for the local variables used in Column (2) of Table 1 along with local and time fixed-effects.

are less of a concern here. Moreover, keeping total employment concentration constant in Column (6) shows that results are specific to high-skill concentration.

We can see the non-monotonicity visually in Figure 4 where I plot on the left-hand side the binned scatter plot between local GDP per-capita growth and the predicted value from the first stage of the 2SLS estimation (i.e. $\widehat{HSConc}_{i,t-1}$). As the threshold $p$ is defined over high-skill concentration level and not over the first-stage predicted values, I also plot on the right-hand side the binned scatter plot where I split the values of $\widehat{HSConc}_{i,t-1}$ between those where high-skill concentration ($HSConc_{i,t-1}$) is below or above the threshold $p$. We can clearly observe the non-monotonic shape, which visually constitutes an inverted-U.

I also assess the importance of sector-level correlation by calculating the exposure-robust standard errors recommended in Adão, Kolesár and Morales (2019). One concern with the "shock-based" identification strategy for SSIVs is that localities with a similar sectoral composition (i.e. similar employment shares $s_{in,t-3}$) may present correlated errors in Equation 3 which are not taken into account when we use municipality-clustered standard errors. Adão et al. (2019) develop "exposure-robust" standard errors which can be extended to a case with an interacted endogenous variable and multiple SSIVs.[32] I calculate these exposure-robust standard errors for specifications in Columns (2) and (3)

[32]I follow one of the author's additional notes on extensions to cases with multiple regressors and instruments.

in Table 1. For Column (2), the exposure-robust errors are 0.356 (5% significance CI = [0.206,1.600]) and 0.161 (5% significance CI = [-0.803,-0.174]) for the bottom and top coefficients, respectively. As for Column (3), the exposure-robust errors are 0.307 (5% significance CI = [0.371,1.574]) and 0.154 (5% significance CI = [-0.801,-0.197]) for the bottom and top coefficients, respectively. Although robust standard errors are larger, coefficients remain significant, as shown by the confidence intervals, and conclusions are unchanged.

Results, then, complement our findings on human capital supply in identifying skill concentration as the underlying channel that is able to explain heterogeneity in growth. We started this analysis by showing evidence that local skill supply only led to higher economic growth in municipalities with low skill concentration among firms. The channel that explains this finding can be summarized as follows. As high-skill supply increases, local large firms benefit relatively more than small firms, increasing skill concentration. The latter, however, has a non-monotonic relationship with growth depending on the level of skill concentration. Hence, an increase in high-skill concentration, due to higher skill supply, causes higher growth in places with low skill concentration and lower growth in places where such concentration is high.

As with the link between skill supply and concentration, we can gauge the economic importance of the link between skill concentration and growth. As shown in Figure A2, high-skill concentration in large firms increased from around 42% in 1999 to around 67% in the 2010s. We can then use our baseline estimates in Column (2) of Table 1 to assess the potential average yearly change in growth rates from the increase in concentration. As with the difference-in-differences estimation, local estimates do not translate easily into the aggregate economy due to the missing intercept problem. Nonetheless, this exercise is useful to gauge whether the high-skill concentration channel is relevant or not. To do so, we assume all municipality-year pairs undergo a 25 percentage-point increase in their local high-skill concentration. We can then calculate the population-weighted average change in growth rates and the long-term percentage-point change in growth. Doing so implies a long-term decline of 1.07 percentage-point, or almost half of per-capita real GDP growth in the period for Brazil.[33] Even if we only consider the increase in skill concentration from the increase in high-skill supply, I estimate a decline in long-term growth of around 18.3%.[34] Notice that here we are assuming that the only effect of an increase in skill supply is to raise skill concentration. As such, the 18.3% decline represents the partial offsetting effect of skill concentration in the link between human capital and growth,

[33] I assess the increase of 25 percentage-points over a period of 11 years. The national per-capita growth rate between 1999 and 2010 was 2.25%.

[34] To arrive at this number, I use the estimated increase in aggregate skill concentration from the aggregate change in high-skill supply using results from Section 3.1.

whereas the total net effect corresponds to the decline shown in Figure A6 of around 6%. Estimates show that the skill concentration channel can more than offset the positive effects of skills on growth.

### 3.3 Shift-Share Design: Skill Premium

We can leverage the SSIV design in Section 3.2 to study the effect of high-skill concentration at large firms on the skill premium. While the latter is not part of my key result linking human capital supply and economic growth, it will be useful as additional validation when I formalize my findings in a model with endogenous growth as the increase in skill concentration can have secondary effects other than on growth.

We proceed by using the same SSIV design as before. Explicitly, I use a similar specification to Equation 3 where I replace GDP per-capita growth as the dependent variable with local skill premium, here defined as the ratio between wages for high-skill and non-high-skill workers at a municipality. I then show estimation results in Table A.7 in the Appendix which follows the same framework as Table 1. In particular, I instrument high-skill concentration with the public loans SSIV described in Section 3.2. However, an issue with this estimation is that the SSIV calculated using loans to small firms gets weaker when high-skill concentration is high. This is expected as small firms play a less significant role in the local economy when concentration at large firms is high. While this is not an issue in itself, when estimating the effect on skill premium this leads to a failed overidentification test due to spurious coefficients from $B_{i,t-2,small}$. As such, I remove $B_{i,t-2,small}\mathbb{1}\{\{HSConc_{i,t-1} > p\} = 1\}$ from the set of instruments. In Section A.4 I show evidence that this is caused by observations with high levels of high-skill concentration and that results are robust to instrumenting with $B_{i,t-2,small}$ without an interaction term.

Results for the skill premium are similar to those for GDP growth. The data shows a non-monotonic relationship between high-skill concentration and the skill premium: while an increase in concentration leads to an increase in the skill premium at low levels of concentration, further increases in high-skill concentration reduce the skill premium. Coefficients are significant in all specifications implying results are robust to controls and potential threats to the identification strategy. Similar to Table 1, joint F-statistics are above the usual weak-IV threshold except for Column (7), and the effective F-statistics calculated individually for each regressor of interest are above the critical values. Finally, we do not reject the null for the J-test of over-identification when using more than one instrument, providing support for identification.[35]

[35] Differently from Table 1 I do not run a LIML specification when using all endogenous regressors since, in this case, we are dealing with a just-identified specification. Hence, both 2SLS and LIML yield identical results.

As in Section 3.2, we can gauge the importance of the high-skill concentration channel to movements in aggregate skill premium. As I show in Figure A26 in the Appendix, there has been an important decline in the skill premium in Brazil. While different factors can affect high and low-skill wages, we can then use our estimates in Column (2) of Table A.7 to assess the relative importance of the high-skill concentration channel. Doing a similar back-of-the-envelope calculation as the one in Section 3.2 yields a 0.29 drop in the skill premium over 11 years from the increase in aggregate skill concentration. This decline is quite significant as it can potentially explain the entire aggregate decline in skill premium between 1999 and 2010. As such, the large increase in high-skill concentration has led to a significant decline in skill premium, as well as in growth.

This novel non-monotonicity result between the skill premium and high-skill labor concentration extends the existing literature on monopsony power in the labor market. As shown in Azar et al. (2022), Schubert et al. (2024), and Jarosch, Nimczik and Sorkin (2024), as firms gain more power in the labor market they push wages down. While this is reasonable regarding unskilled labor, my results show that high-skill wages react differently to labor market power given the R&D competition dynamics. I rationalize this finding in Section 4 by combining in a single model both labor search and step-by-step innovation.

### 3.4 SSIV Falsification Tests

Next, I assess the validity of the shift-share instruments and the shift-approach through falsification tests. While it is encouraging that we did not reject the null hypothesis in the overidentification tests in Table 1, I also test the shock exogeneity assumption using the tests proposed in Borusyak et al. (2021). Naturally, falsification tests cannot prove instrument validity. Nonetheless, passing results strengthen the identification assumption.

I start by showing how much variation we have at the shock level. This is important as the validity assumption requires enough variation at the shock level for consistency. Table A.8 in the Appendix shows the summary statistics for both shocks $g_{n,t-2}$ and shares $s_{in,t-3}$ in my sample, split between large and small firms. For the shocks, statistics are weighted by the shares and I also report statistics on the distribution of shocks after residualizing with year fixed-effects (weighting with shares). That is, I regress shocks on year fixed-effects while weighting with shares. This allows us to gauge whether there is enough within-period variation. In addition, I report, as suggested by Borusyak et al. (2021), the effective sample size measured as the inverse of the share Herfindahl-Hirschman index, i.e. $1/\sum_{n,t} s_{n,t-3}$ where $s_{n,t-3} = \sum_i s_{in,t-3}$ are the sector-level shares. This HHI is a measure of how concentrated sector exposure is and, hence, measure whether we have enough sector-level variation for asymptotic validity. Borusyak et al. (2021) show through Monte Carlo simulations that an effective sample of at least 20 provides enough variation for large-sample approximations at the shock level.

Table A.8 shows that we have sizeable variation at the sector level. Largest shares for both small and large firms are 1.1% and 5.1% respectively, indicating that no single sector-period has an overweight on the distribution. Shock distributions for both large and small firms look regular and have standard deviations that are larger than their means. Residualizing shock distributions for large and small firms with year fixed-effects only has a significant effect on the former as the standard deviation drops by around 50%. However, the effective sample for both large (26) and small (190) local-level shares are above the threshold of 20 which suggests high enough variation.

I then implement falsification tests at the shock level. These tests consist of regressing sector-level controls and the lagged dependent variable (i.e. GDP growth per capita), both taken prior to the realization of shocks $g_{n,t-2}$, on shocks directly weighting by the shares. Formally, let $q_{it}$ be a control variable used in Equation 3. We then run:

$$\overline{q}_{n,t-3} = \beta g_{n,t-2} + \gamma V_{n,t-2} + \epsilon_{n,t} \tag{6}$$

where $\overline{q}_{n,t-3} = \frac{\sum_i s_{int} q_{i,t-3}}{\sum_i s_{int}}$ is the exposure-weighted average of $q_{i,t-3}$ and $V_{n,t-2}$ is the set of all controls used in Equation 3, including time fixed-effects, except $q_{i,t-2}$.[36] When using the lagged dependent variable on the left-hand side, I replace $q_{i,t-3}$ with $y_{i,t-3}$ (i.e. GDP per capita growth or skill premium). Finally, I also use data from sector-level surveys at the national level to check for balance between sectors on supply-side parameters, although I can only run this specification on the combined shock to both small and large firms (vs. running it separately for small and large-firm shocks).

The intuition behind this test is two-fold. First, in assessing whether there are any significant correlations between shocks and prior observables at the sector level we can look for significant differences between industries exposed to large shocks relative to those under small shocks. If we find any, we may potentially worry that our results in Table 1 and Table A.7 are biased due to correlations with unobservables even though I control for the observables being tested (Oster, 2019). Second, regressing lagged GDP growth and skill premium on shocks provides us with a pre-trend test similar to difference-in-differences specifications. A significant shock coefficient could indicate that high-shock sectors were on a different trend relative to low-shock sectors prior to the realization of the shock, posing a threat to identification.

[36] I run the regression at the sector level in order to avoid the clustering issue shown in Adão et al. (2019).

Table 2: Shock balance tests and pre-trend tests

| | GDP Growth | | Skill Premium | | Log(Wage) | | Log(Population) | | % High-Skill | | % Min. Wage | | Net Hiring | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| Shock - Large Firms | 0.0942 | | 0.134 | | 0.171 | | 0.0145 | | -0.00764 | | -0.0382* | | 0.0198 | |
| | (0.061) | | (0.072) | | (0.092) | | (0.025) | | (0.024) | | (0.019) | | (0.026) | |
| Shock - Small Firms | | -0.0560 | | 0.0130 | | -0.0189 | | -0.0317 | | 0.0379 | | -0.00920 | | 0.000947 |
| | | (0.047) | | (0.077) | | (0.095) | | (0.055) | | (0.019) | | (0.016) | | (0.026) |
| N | 593 | 622 | 593 | 622 | 593 | 622 | 593 | 622 | 593 | 622 | 593 | 622 | 593 | 622 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Dependent variable is defined as $\overline{q}_{n,t-3} = \frac{\sum_i s_{int} q_{i,t-3}}{\sum_i s_{int}}$ where $s_{int}$ are the exposure shares and $q_{i,t-3}$ is one of the controls used in Equation 3. Regressions are weighted by sector high-skill employment shares. In Columns (1)-(4), $q_{i,t-3}$ is replaced by $y_{i,t-3}$ where $y_{i,t-3}$ is GDP per-capita growth in Columns (1)-(2) and skill premium in Columns (3)-(4), both winsorized at the $1^{st}$ and $99^{th}$ percentiles. Local-level controls (all lagged to be contemporaneous with shocks): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. In each specification the variable used as the dependent variable is excluded from the list of controls. Sector-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

I show results for the falsification and pre-trend tests in Table 2 for both large and small firm shocks using local level variables, and in Table 3 using sector survey data. In the former, Columns (1)-(4) show the pre-trend tests, both for local growth rates and the skill premium, while Columns (5)-(14) show the local balance test. All variables have been demeaned and normalized to have unit variance so that coefficients are more easily interpretable, and standard errors are clustered at the sector level. All but one coefficient pass the balance test of non-significant results, i.e. we do not observe shock imbalance with respect to wages, population size, the percentage of high-skill workers in the population, and the population ratio of net hiring. Moreover, we observe no pre-trends with respect to GDP growth and skill premium. While the shock coefficient for large firms when regressing the population share of workers receiving minimum wage or less is statistically significant, the magnitude is small: a one standard deviation increase in the shock is associated with a decline in the minimum wage share by approximately -4% of its standard deviation. This difference between high and low-shock sectors does not seem sensitive enough to drive results.

As for the sector survey data, I find no significant shock imbalance with respect to supply-side variables. These consist of the growth in net revenues and value added, the ratio of wages, intermediate inputs costs, and fuel and electricity costs to value added, and the share of production workers over total sector employment, either measured at the end of the year or as an yearly average. While sector survey coverage is lower than the one in the RAIS database, I manage to cover most sectors. All coefficients between

Table 3: Sector-level survey data balance check

| Balance Variable | Coef | SE | Obs. |
|---|---|---|---|
| Revenue Growth | -0.161 | (0.134) | 495 |
| Value Added Growth | -0.026 | (0.028) | 480 |
| Wages-to-Value Added Ratio | 0.041 | (0.046) | 495 |
| Intermediate Inputs-to-Value Added Ratio | -0.014 | (0.033) | 495 |
| Fuel and Electricity-to-Value Added Ratio | 0.039 | (0.089) | 473 |
| Production Workers' Share of Employment (on 12/31) | 0.020 | (0.050) | 359 |
| Production Workers' Share of Employment (yearly avg.) | 0.054 | (0.055) | 359 |

Table reports the regression coefficients of each sector-level variable on shocks $g_{n,t}$ weighted by each sector's high-skill employment share and controlling for year fixed-effects. Variables are set to the shocks' initial period ($t-1$). Standard errors are sector-clustered. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

variables and shocks are not significant, a result we would expect if shocks are indeed as-good-as-randomly assigned to industries each year. [37]

With enough shock-level variation and having passed the falsification tests, the *a priori* assumption of shock exogeneity for my SSIV seems plausible. Although local demand for loans from the Brazilian national development bank depend on local conditions, changes to the national amount disbursed to different sectors and firm sizes seem exogenous to municipality-level conditions, and we fail to reject imbalance between sectors. The evidence, then, points to the validity of the SSIV identification strategy.

Finally, I leave to Section A.4 in the Appendix a series of robustness checks of the SSIV results. First, I show that the non-monotonic results are robust to using polynomial regressors and instruments instead of interacting both with a threshold $\mathbb{1}\{HSConc_{i,t-1} > p\}$. Second, results remain unchanged if we use a narrower definition of high-skill workers which includes information on the type of skills required for different occupations. Third, I show that the overidentification test in the skill premium specification fails due to heterogeneity in the slope between skill premium and skill concentration. Fourth, results are robust to restricting the sample to the non-tradable sector only, which is reassuring as my skill concentration mechanism involves competition in local labor markets. Finally, I show robustness to several additional changes to the specification, including running a weighted regression weighting by the log of local population, lagging the SSIV exposure shares one additional period, and to changes in the threshold $p$ that defines places where skill concentration is high or low.

[37] Another concern when using SSIVs is that a strong serial correlation of shocks, combined with latent dynamic adjustments of the dependent variable in Equation 3, may bias our results (Jaeger, Ruist and Stuhler, 2018). In our case, $g_{n,t-2,large}$ and $g_{n,t-2,small}$ have a serial correlation of -0.047 and -0.083, respectively. As such, any dynamic bias would not affect results significantly.

# 4 Model

I now rationalize my findings from Section 3 in an endogenous growth model with high-skill labor demand and search. I first describe the model's framework. I then show GMM estimation results and how they relate to the empirical findings in the previous section.

## 4.1 Model Framework

We start with a closed economy in continuous time and a unit continuum of markets $j$ where two firms compete in a technology ladder in each market (similar to Aghion et al., 2001, Acemoglu and Akcigit, 2012, Liu et al., 2022).[38] In each market $j$ there is also a firm producing a non-innovative good, i.e. there are two goods in each market: one produced by the competing R&D firms and one produced by the non-innovative firm, referred to as $i$ (or $-i$) and $o$ respectively. At any moment in time an innovative firm $i$ is located at step $m$ of the technology ladder. Consumers have log-utility preferences over consumption, own firms in the economy, and provide one unit of work of one out of two types: high or low skill. Intertemporal preferences are as follows:

$$
\begin{aligned}
&U = \int_0^\infty e^{-rt}\Big\{\int_0^1 \nu \ln x_j(t) + (1-\nu)\ln x_{o,j}(t)dj\Big\}dt \\
&s.t. \int_0^1 p_{i,j}(t)x_{i,j}(t) + p_{-i,j}(t)x_{-i,j}(t) + p_{o,j}(t)x_{o,j}(t)dj = w_k(t)l_k(t) + \pi(t) \\
&x_j(t) = x_{i,j}(t) + x_{-i,j}(t)
\end{aligned}
\tag{7}
$$

where $x_{i,j}(t)$ is demand for firm $i$'s product $j$, $p_{i,j}(t)$ is the price of firm $i$'s product $j$, $w_k(t)$ is the wage ($k = H, L$, high or low skill respectively), $l_k(t)$ is labor, $\pi(t)$ are profits, $\nu \in (0,1)$, and $r$ is the discount rate.

Innovative firms $i$ engage in two activities: production and research. While innovation requires high-skill workers, production uses low-skill ones. The production function for firm $i$ follows:

$$y_i(t) = \gamma_i(t)l_{i,L}(t) \tag{8}$$

where $y_i(t)$ is output, $\gamma_i(t)$ is productivity, and $l_{i,L}(t)$ is low-skill labor. Productivity

[38] While these models are usually applied to developed economies, there is relatively less frontier R&D effort in Brazil. Nonetheless, innovation models can still offer useful insights if we consider a broader definition of both "innovation" and "R&D" which includes process innovation and adopting foreign technologies (vs. catching up to a domestic market leader).

evolves according to the following law of motion:

$$\gamma_i(t+\Delta t) = \begin{cases} \gamma^{m+1}, \ if \ R\&D \ successful \\ \gamma^{m}, \ if \ R\&D \ fails \end{cases} \tag{9}$$

where $\gamma > 1$ is a constant. Equation 9 implies that each successful R&D effort moves the firm one step further in the technology ladder. The arrival rate of successful innovation happens at a Poisson rate $\eta_i(t)$ which is determined by the following R&D production function:

$$\eta_i(t) = A_\lambda \lambda_i(t) + A_l l_{i,H}(t)^\alpha \tag{10}$$

where $A_\lambda$ and $A_l$ are constants, $\lambda_i(t)$ is R&D investment, $l_{i,H}(t)$ is high-skill labor, and $\alpha \in (0,1)$ is a constant. We assume low-skill labor supply is perfectly elastic and paid at an exogenous wage $w_L$, while high-skill labor supply is paid $w_H$ which is determined through labor search.[39] The cost of investing $\lambda_i(t)$ is quadratic, i.e.:

$$C(\lambda_i(t)) = \rho \frac{\lambda_i(t)^2}{2} \tag{11}$$

where $\rho$ is a constant.

Innovative firms compete a la Bertrand.[40] Define the technological gap between two firms in a market as $s(t) = m_i(t) - m_{-i}(t)$. We shall call the firm that is ahead the "leader" (henceforth, referred by the subscript L) and the one that is behind the "follower" or "laggard" (subscript F). As such, for $s > 0$ the leader takes the whole market and charges a price that is the marginal cost of its competitor. For $s = 0$, both firms split the market equally. Then, from log-utility:

$$x_L(t) = \frac{\nu D(t)}{p_{i,j}(t)}\ , \ x_F(t) = 0 \tag{12}$$

where $D(t) = w_L(t)l_{i,L}(t) + w_o(t)l_o(t) + \pi(t)$ is aggregate demand.[41] It is straightforward to show that the optimal low-skill labor demand for the leader when $s > 0$ is:

$$l_{i,L}(t) = \frac{\nu D}{\gamma^s w_L(t)} \tag{13}$$

[39] I assess the assumption of perfect elasticity of low-skill labor supply in Sections A.5 and A.6 in the Appendix. I also assess results under different labor market assumptions in Section A.6.

[40] Model results under an assumption of competition a la Cournot remain qualitatively similar.

[41] Since high-skill wages are paid out of profits, only $\pi(t)$ shows up.

We can then write the static problem for the innovative firms. Consider a leading firm $i$ who is $s$ steps ahead from the laggard. Profits can be written as (I henceforth drop the time dependency to simplify the notation):

$$\pi_s = \max_{p_{i,j}} \left(p_{i,j} - \frac{w_L}{\gamma^{m+s}}\right)x_{i,j} = \left(\frac{w_L}{\gamma^m} - \frac{w_L}{\gamma^{m+s}}\right)\frac{\nu D_s}{p_{i,j}} = (1-\gamma^{-s})\nu D_s \tag{14}$$

Given Bertrand competition follower's profits are zero, i.e. $\pi_{-s} = 0$. When $s = 0$ the industry is "neck-and-neck" and both firms make no profits ($\pi_0 = 0$). Firms decide strategically on how much to invest in R&D ($\lambda_i$) and how much high-skill labor to hire ($l_{i,H}$) as they have to consider the technological gap $s$ with their competitor. Conditional on $s$, profits are no longer time-dependent nor do they depend on where each firm is on the technology ladder.

Regarding the non-innovative firm, it only engages in production via the same production function as in Equation 8. However, differently from the R&D firms it employs high-skill labor in production, i.e. $w_o = w_{o,H,s}$ and $l_o = l_{o,H,s}$. This aspect of the model captures an important fact about the Brazilian economy which is that a significant share of high-skill workers does not work in jobs that require a college degree (38% in 2018, Lameiras and Vasconcelos, 2018).[42] Since such employees still earn more than low-skill workers, we assume that the non-innovative firm has to hire its workers through search. I show later on when estimating the model that adding a non-innovative firm that hires high-skill labor does not affect the qualitative results regarding the innovative firms, though it will prove important quantitatively to match labor market empirical moments. To guarantee the existence of a balanced growth path, we assume the productivity of the non-innovative firm $\gamma_o$ grows at the same rate as the expected growth rate of $\gamma_s$.

High-skill labor search works similarly to the Diamond-Mortensen-Pissarides (DMP) framework (Diamond, 1982, Mortensen, 1982, Pissarides, 1985) where high-skill workers are either employed in R&D or searching for work while being unemployed. One important difference relative to the DMP framework is that I make an assumption, explained below, that removes the necessity of keeping track of a firm's current labor force. Let $u_s$ be the unemployment rate when the gap between both innovative firms is $s$ and $v_s$ ($v_{-s}$, $v_{o,s}$) the vacancies posted by the leader (follower, non-innovative firm) such that $\overline{v}_s = v_s + v_{-s} + v_{o,s}$. Let the matching function $M$ be defined as:

$$M(u_s, \overline{v}_s) = B u_s^{\varphi} \overline{v}_s^{1-\varphi} \tag{15}$$

[42]Realistically, non-innovative firms may hire both high and low-skill workers for production, potentially with different labor productivities. As allowing both types of labor would not change anything significantly in the model, we make the simplifying assumption that such firms only hire high-skill labor.

where $\varphi$ is a constant. Define $\theta_s \equiv \overline{v}_s / u_s$ as the labor market tightness. Then the worker flow rate from unemployment to employment is $M/u_s = B\theta_s^{1-\varphi}$ and the vacancy matching rate for a firm posting $v_s$ vacancies is $v_s M/\overline{v}_s = v_s B\theta_s^{-\varphi}$.

Let the cost for a firm of posting a vacancy be:

$$C_{v,s} = \kappa \frac{v_s^2}{2} \tag{16}$$

where $\kappa$ is a constant.

We can now define the value functions for high-skill workers and firms. Let $W_s$ be the value of employment and $U_s$ be the value of unemployment for a worker. The value function of being a high-skill worker is:

$$rW_s = w_{H,s} + \delta(U_s - W_s) \tag{17}$$

where $r$ is the interest rate and $\delta$ is an exogenous separation constant. Equation 17 is straightforward: while employed at a leading firm $s$ steps ahead, a high-skill worker receives wage $w_{H,s}$ and faces an exogenous probability of being laid-off.

Conversely, the value of unemployment is:

$$rU_s = b + \frac{v_s}{\overline{v}_s} B\theta_s^{1-\phi}(W_s - U_s) + \frac{v_{-s}}{\overline{v}_s} B\theta_s^{1-\phi}(W_{-s} - U_s) + \frac{v_{o,s}}{\overline{v}_s} B\theta_s^{1-\phi}(W_{o,s} - U_s) \tag{18}$$

where $b$ is the value of the outside option. As with Equation 17, Equation 18 describes the change in value flow for an unemployed worker who can find a vacancy from either firm $s$, $-s$, or $o$.

I then make an important change regarding innovation diffusion relative to previous models of strategic interaction. In the original set-up (Aghion et al., 2001), the follower pays for an arrival rate of innovation of $\eta_{-s}$ yet gets $\eta_{-s} + h$, where $h \geq 0$ is a constant that represents the relative easiness of catching up to the leader.[43] Instead, I consider the case where the diffusion parameter is a function of the high-skill labor currently working at the laggard firm. As such, the follower gets $\eta_{-s} + (h_l l_{-s,H}^{\alpha} + h_c)$, where $h_c, h_l \geq 0$ are constants. To simplify, I assume high-skill workers at the follower firm work in internal and catch-up R&D at the same time. I consider the case of separate hiring and employment when analyzing counterfactuals.

The idea behind making the innovation catch-up a function of high-skill labor is two-fold. First, it brings the model closer to reality as firms have to develop internal capacity in

[43] Although this term is referred to as an "imitation" parameter in Aghion et al. (2001), I interpret it here as a parameter that regulates the diffusion of ideas from the innovation frontier to the firm catching up.

order to absorb external knowledge even if such knowledge can be seen as a public good. Second, it strengthens the link between the R&D efforts of leader and laggard through the labor market. As the leader hires skilled labor, the labor market becomes tighter allowing the leader to indirectly hinder the innovation catch-up efforts of the laggard. Results will show a decline of active R&D catch up, that is the catch-up effort due to a firm's own high-skill hiring, with an increase in skill concentration.

We can, then, write the dynamic problem of the innovative firms as a function of the R&D gap $s$:

$$\begin{aligned} rJ_s = \max_{\lambda_s, l_{s,H}} \Big\{ & \pi_s - \rho\frac{\lambda_s^2}{2} - w_{s,H} l_{s,H} - \kappa\frac{v_s^2}{2} + [A_\lambda \lambda_{-s} + A_l l_{-s,H}^{\alpha} + h_l l_{-s,H}^{\alpha} \\ & + h_c](J_{s-1} - J_s) + [A_\lambda \lambda_s + A_l l_{s,H}^{\alpha}](J_{s+1} - J_s) \Big\} \end{aligned} \tag{19}$$

$$\begin{aligned} rJ_{-s} = \max_{\lambda_{-s}, l_{-s,H}} \Big\{ & \pi_{-s} - \rho\frac{\lambda_{-s}^2}{2} - w_{-s,H} l_{-s,H} - \kappa\frac{v_{-s}^2}{2} + [A_\lambda \lambda_s + A_l l_{s,H}^{\alpha}](J_{-s-1} - J_{-s}) \\ & + [A_\lambda \lambda_{-s} + A_l l_{-s,H}^{\alpha} + h_l l_{-s,H}^{\alpha} + h_c](J_{-s+1} - J_{-s}) \Big\} \end{aligned} \tag{20}$$

$$\begin{aligned} rJ_0 = \max_{\lambda_0, l_{0,H}} \Big\{ & \pi_0 - \rho\frac{\lambda_0^2}{2} - w_{0,H} l_{0,H} - \kappa\frac{v_0^2}{2} + [A_\lambda \lambda_{-0} + A_l l_{-0,H}^{\alpha}](J_{-1} - J_0) + \\ & [A_\lambda \lambda_0 + A_l l_{0,H}^{\alpha}](J_1 - J_0) \Big\} \end{aligned} \tag{21}$$

where $(\lambda_{-0}, l_{H,-0})$ refers to the competing firm at $s = 0$.

The dynamic problem in Equation 19 can be understood as follows. For the leader (first and second lines), it receives a static flow of profits and has to pay the innovation investment cost, the high-skill labor wage, and the cost of posting vacancies. At a rate $\eta_{-s} + h_l l_{-s,H}^{\alpha} + h_c$ the follower is able to reduce the gap relative to the leader from $s$ to $s-1$. Conversely, the leader is able to increase the gap by one at a rate $\eta_s$. The situation is analogous for the follower and neck-and-neck firms.

At this point, I make one important simplifying assumption. Both the firm's dynamic problem in Equation 19 and the high-skill labor search problem, which yields the high-skill wage, need to be solved simultaneously as both require the firm's value function $J_s$ for all $s$. Moreover, in the usual search model framework labor is a state variable, i.e. we have to keep track of how many workers a firm currently has and solve the problem at every value of the gap $s$. To make things tractable, I split the firm's decision into two steps. First, the firm searches for high-skill labor until it hires the optimum amount $l_{H,s}^*$ for its current gap $s$. Then, it engages in R&D and finds out whether it was successful or not (along with its competitor). Another way of understanding this simplification is to assume that the firm hires through *collective hire bargaining*: it gathers all the high-skill workers it found and makes a collective offer to hire all of them at once. This assumption can also be understood from a time-frame perspective: by the time a firm successfully innovates, it has already managed to hire the amount of labor it wants given $s$, i.e. labor adjusts quicker relative to the time between two innovation steps.[44] As a result, labor is no longer a state variable and we only need the value of labor demand at the steady state for each $s$.

This simplifying assumption, which effectively implies that firms achieve their desired level of labor demand before engaging in R&D, allows us to get an equation for high-skill labor demand in steady-state where $l_{s,H}(t) = l_{s,H}(t+1) = l^*_{s,H}$:

$$l^*_{s,H} = (1-\delta)l^*_{s,H} + v_s B\theta^{-\varphi} u_s \tag{22}$$

I provide the derivation for the choice of optimal investment $\lambda_s$ and labor demand $l_{s,H}$ in Section A.5 in the Appendix.

As for the non-innovative firm, it solves the following static problem every period:

$$\pi_{o,s} = \max_{l_{o,H,s}} p_{o,s}\gamma_o l_{o,H,s} - w_{o,H,s} l_{o,H,s} - \kappa \frac{v^2_{o,s}}{2} - c_{f,s} \tag{23}$$

where $c_{f,s}$ is a fixed cost which we add to make $\pi_{o,s} = 0,\ \forall s$ without loss of generality. This not only simplifies the wage equation later on but also highlights how results about the R&D firms will not depend on the non-innovative sector. Analogous to the R&D sector, demand for the non-innovative good is $y_{o,s} = (1-\nu)D_s/p_{o,s}$. We can then solve Equation 23 using Equation 22 to get labor demand at the non-innovative firm.

The final step is to solve the labor search problem. I define the net value of a match (i.e. the surplus) as follows:

$$S_s \equiv W_s - U_s + J_s - V_s \tag{24}$$

where $V_s$ is the value function of the firm when it hires no labor, i.e. when collective hire bargaining has failed.[45] To solve the bargaining problem between firm and workers, I adopt the usual Nash bargaining solution. Let $\xi$ be the weight for workers. We can, then, write the surplus as:

$$\xi S_s = W_s - U_s \tag{25}$$

[44] I, hence, assume the transitory effect on R&D effort from adjusting labor between $l_{H,s}$ to $l_{H,s'}$ to be of second order.

[45] To get $V_s$, we have to solve a version of Equation 19 where collective hiring fails. For simplicity, I assume that firms do not invest in R&D when collective hiring fails (though they may do so if labor demand is zero) and pay the same vacancy costs as if hiring was successful.

$$(1-\xi)S_s = J_s - V_s \tag{26}$$

By plugging-in Equation 26 into Equation 25 along with the definitions of $W_s$ and $U_s$ in Equations 17 and 18, we arrive at the following expression for high-skill wage at the leading R&D firm:[46]

$$w_{s,H} = b + \xi S_s(r+\delta) + \xi B\theta_s^{1-\varphi}\left[\frac{v_s}{\overline{v}_s}S_s + \frac{v_{-s}}{\overline{v}_s}S_{-s}\right] \tag{27}$$

Finally, we require the following labor market clearing conditions:

$$\begin{aligned} L_H &= l_{s,H} + l_{-s,H} + l_{o,H,s} + u_s L_H \\ L_L &= l_{s,L} + l_{-s,L} \end{aligned} \tag{28}$$

where $L_H$ ($L_L$) is the total amount of high-skill (low-skill) labor that is available locally.

## 4.2 Model Estimation

We can now solve for the steady state. This requires us to pin-down 15 parameters: $\{\xi, \varphi, \delta, \alpha, r, B, \gamma, b, \rho, A_l, A_\lambda, \kappa, h_l, h_c, \nu\}$. First, I set $L_H = 1$ and the low-skill wage $w_L$ to match the in-sample average which is R\$1,1734.8 monthly.[47] I then pick $\xi = 0.45$ for the bargaining power parameter following Ulyssea (2010), which is close to the usual value in the literature (0.5). I set the elasticity with respect to unemployment $\varphi$ in the matching function to 0.5 (Petrongolo and Pissarides, 2001, Ulyssea, 2010, Dix-Carneiro, Goldberg, Meghir and Ulyssea, 2021). I calculate the separation rate for high-skill workers in sample and set $\delta = 0.084$ which is the average among municipality-year pairs. $\alpha$ is set to 0.438 in line with the estimate in Growiec, McAdam and Muck (2023) for a TFP production function. I calibrate $r$ to the average nominal baseline interest rate (SELIC) deflated with the 12-month inflation expectation series for the period between 2000 and 2017. This gets us $r = 8\%$.

As for the matching function scaling parameter $B$, I calibrate it to the following unemployment flow equation which equates flows from and to unemployment:

$$\delta(L_H - E[u_s]L_H) = BE[\theta_s]^{1-\varphi}E[u_s]L_H \tag{29}$$

where $E[.]$ is the expectation operator. We then set $E[u_s] = 6.07\%$ and $E[\theta_s] = 0.48$ to

[46] I provide the formal proof of Equation 27 in the Appendix. Notice from the $\pi_{o,s} = 0$ condition that the surplus for the non-innovative firm is zero.

[47] To get the annual wage, I multiply the monthly rate by 13 to take into account the mandatory end-of-the-year bonus which is equivalent to a month's payment.

arrive at $B = 1.88$.[48]

That leaves us with nine remaining parameters to estimate: $\{\gamma, b, \rho, A_l, A_\lambda, \kappa, h_l, h_c, \nu\}$. I do so via a GMM estimation using the following 10 empirical moments: average real GDP per capita growth rate, average skill premium at large firms weighted by number of workers, average labor market tightness, average high-skill wage at non-large firms weighted by number of workers, average high-skill labor concentration at large firms, average firm profitability, R&D share of sales, average cost of hiring per job, average unemployment of high-skill workers, and share of markets where high-skill concentration is below or equal to 50%. While there is no 1:1 mapping between parameters and moments, especially since moment fit depends on the distribution of sectors over gaps $s$, we can associate sets of parameters to their most closely related moments. R&D investment cost parameter $\rho$ directly influences the R&D investment-to-sales ratio. Similarly, we can pin-down the vacancy cost scalar $\kappa$ with the average cost of hiring. Firm profitability only depends on $\gamma$. $\nu$ influences labor market tightness and high-skill unemployment as the non-innovative sector hires most of the labor supply. These moments, along with the skill premium, are also influenced by the value of the outside option $b$ and R&D labor productivity $A_l$. Finally, $h_l$ and $h_c$ help us pin-down high-skill labor concentration, both on average and the sector distribution.

It remains to derive the expression of the growth rate in the model. Note that in steady state both leaders' and followers' productivities grow at the same rate $g$ while the average gap $s$ remains the same. As R&D follows a Poisson arrival, leader productivity improves, in expectation, by $\gamma\eta_s\Delta t$ while the follower's improves by $\gamma[\eta_{-s} + (h_l l^\alpha_{-s,H} + h_c)]\Delta t$. Under such steady state, the inflow and outflow of firms between gap levels $s$ have to balance. Let $\mu_s$ be the share of sectors where the gap between leader and follower is $s$. Then:

$$\begin{aligned} 2\mu_0\eta_0 &= \eta_{-1} + h_l l^\alpha_{-1,H} + h_c \\ \mu_s\eta_s &= \eta_{-(s+1)} + h_l l^\alpha_{-(s+1),H} + h_c,\ s > 0 \end{aligned} \tag{30}$$

where $\sum_s \mu_s = 1$. As such, if we now consider a single sector where a leader and a follower compete, growth can be expressed as:

$$\begin{aligned} g_s &= ln(\gamma)2\eta_0,\ s = 0 \\ g_s &= ln(\gamma)\eta_s,\ s > 0 \end{aligned} \tag{31}$$

while aggregate growth is simply $g_{agg} = \sum g_s\mu_s$. I provide the formal proof of Equation 31 in Section A.5 in the Appendix.

[48] C.f. Section A.7 in the Appendix for details on data moments.

It is worth explaining at this point how I calculate high-skill labor concentration in the model. Importantly, not all high-skill workers are employed at innovative firms, a fact reflected in the data. Yet, for simplicity, we did not split the non-innovative sector between large and non-large firms. However, we will do so now to calculate high-skill concentration. We assume that the high-skill labor in the non-innovative sector is split endogenously between large and small firms according to a Cournot profit split determined by the productivity levels of both the innovative leader and follower firm. Specifically, if two firms compete a la Cournot in the non-innovative sector, one with productivity $\gamma^{m+s}$ and one with productivity $\gamma^{m}$, then it is straightforward to show that profits for both large and small firms can be written as:

$$\pi_{o,s} = \left(\frac{\gamma^{s}}{1+\gamma^{s}}\right)^{2}, \; \pi_{o,-s} = \left(\frac{\gamma^{-s}}{1+\gamma^{-s}}\right)^{2} \tag{32}$$

We can, then, use Equation 32 to calculate high-skill concentration assuming that the large firm share of the non-innovative sector is the large firm profit share, i.e. $\pi_{o,s}/(\pi_{o,s} + \pi_{o,-s})$. I define high-skill concentration in the model as the ratio between employees at large firms, both innovative ($l_{s,H}$) and non-innovative ($\pi_{o,share} l_{o,H,s}$), over the total number of high-skill workers ($l_{s,H} + l_{-s,H} + l_{o,H,s}$).[49] Note, however, that high-skill labor concentration is at its lowest at $s = 0$ since leader and follower are competing neck-and-neck which implies a minimum model-generated level of 50%. This makes model-fit difficult as in reality we observe many municipality-year pairs where concentration is below 50%. As a solution, I first calculate high-skill labor concentration as aforementioned (call it $LC_1$). I then calculate a second measure ($LC_2$) which takes the value of 1 whenever the follower does not hire ($l_{-s,H} = 0$), is a linear function of the gap $s$ when $l_{-s,H} > 0$, and at $s = 0$ we assume $LC_2 = 1/s_{min}$, where $s_{min}$ is the lowest value of $s$ where $l_{-s,H} = 0$. This second measure is more in line with the fact that in reality competition through a quality ladder involves several firms, and that at a neck-and-neck state the interaction between firms looks more like perfect competition. Finally, the model-generated high-skill labor concentration is the average between $LC_1$ and $LC_2$. I show further below that results remain unchanged using different approaches to calculating concentration. Importantly, this only affects model fit as it only affects how we calculate high-skill concentration.

I show GMM estimation results in Table 4. Overall, model fit is good as data and model-generated moments are close, especially for the growth rate and the skill premium at large firms. I further assess the model fit by checking the match relative to a non-

[49] I later show robustness of results if, instead, we ignored the non-innovative sector in calculating high-skill concentration.

targeted moment, i.e. the R&D worker share. Though the non-targeted fit is worse than the targeted ones, it is reassuring that the model-generated value is not far from the empirical moment.

Table 4: Model estimation and moment fit

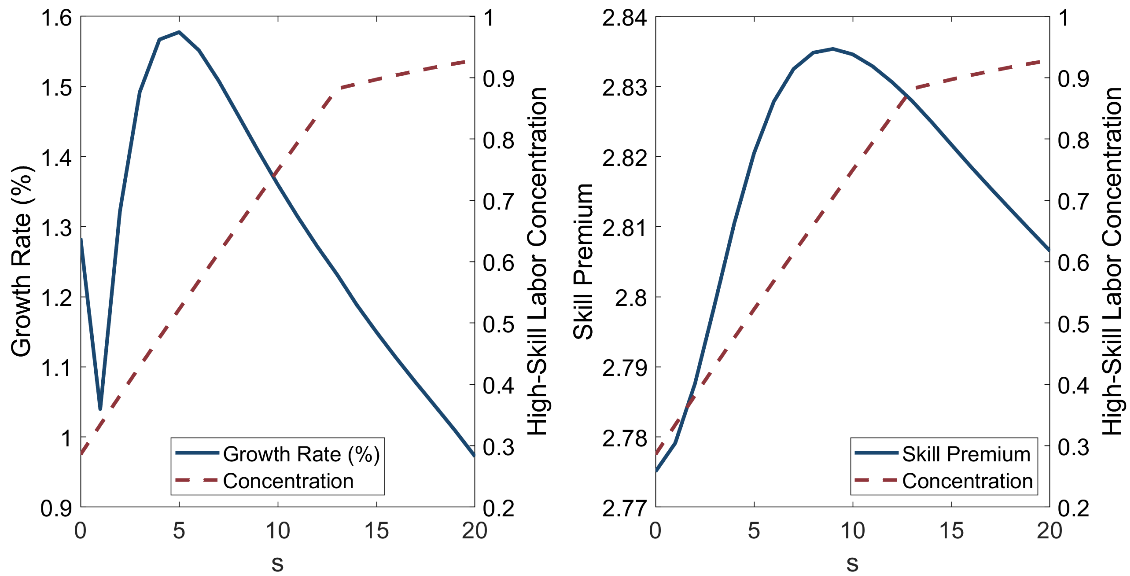
| Parameter | Value | Parameter | Value |
|---|---|---|---|
| $\gamma$ | 1.046 | $\kappa$ | 0.44 |
| $b$ | 0.62 | $h_l$ | 1.7 |
| $\rho$ | 3084.1 | $h_c$ | 0.31 |
| $A_l$ | 2.23 | $\nu$ | 0.21 |
| $A_\lambda$ | 29.6 | | |
| **Moments** | | **Data** | **Model** |
| Growth Rate (%) | | 1.31 | 1.30 |
| Skill Premium, Large Firms | | 2.76 | 2.77 |
| Labor Market Tightness | | 0.48 | 0.48 |
| High-Skill Wage, Non-Large Firms | | 0.58 | 0.51 |
| High-Skill Concentration | | 0.59 | 0.59 |
| Firm Profitability | | 0.20 | 0.24 |
| R&D Investing-to-Sales Ratio (%) | | 0.19 | 0.21 |
| Cost-per-Hire | | 0.045 | 0.037 |
| High-Skill Unemployment | | 0.19 | 0.22 |
| Share of High-Skill Concentration $\leq$ 50% | | 0.38 | 0.42 |
| **Non-Targeted Moment** | | **Data** | **Model** |
| R&D Worker Share (%) | | 0.91 | 0.68 |

We can then analyze the firm's problem and choice. I show in Figure A27 in the Appendix the value function of both leader and follower (left-hand side), and the hiring and investment decisions (right-hand side) as a function of the gap $s$. Starting with the value functions, they are both monotonic: increasing for the leader and decreasing for the follower. At a high-enough $s$ the follower's value function is essentially zero while the leader's value function changes concavity. This change in concavity identifies the region of most intense innovation effort by the leader as it attempts to escape competition from the follower. This can be seen in the right-hand side plot which shows a peak in both R&D investment and high-skill hiring. This is followed by a reduction in R&D effort in the "lazy monopolist" region where R&D effort falters due to the discouragement effect as the gap is too large for any credible competitive threat. These results are expected in step-by-step models of innovation.

The novelty lies in what we gain by adding high-skill labor to the model. This can be seen in Figure 5 where I show the growth rate $g_s$, skill premium,[50] and high-skill labor

[50] The skill premium shown in Figure 5 is calculated only for the innovative firms, i.e. we are ignoring the non-innovative firm. While it is important to take the non-innovative sector into account when matching moments, here I want to highlight the firm interaction in the innovative sector. Results adding the non-innovative firm have a similar inverted-U shape.

Figure 5: *Left: Growth and high-skill labor concentration; Right: Skill premium and high-skill labor concentration, all as a function of the gap s*




concentration as a function of the gap $s$. Except for the neck-and-neck ($s = 0$) region, both plots show non-monotonic curves for the growth rate and the skill premium resembling an inverted-U shape while high-skill labor concentration increases. This captures the same patterns estimated in Sections 3.2 and 3.3 in reduced-form: as concentration increases, at first both the local growth rate and skill premium increase. However, as labor concentration keeps increasing the relationship inverts as growth and skill premium go down. The changes in the growth rate are significant as it moves from around 1.6% at peak to a bottom near 1%. Notice that the reduced-form results capture differences at the municipality level, where each local area is at a different gap $s$. This is why we are analyzing both growth and the skill premium as a function of the gap.

Figure A27 also highlights the importance of search frictions. We can see this clearly if we start form a model with only R&D investment. As firms' incentives to innovate decrease once a leading firm is far ahead, they can adjust investment down accordingly. Once we add high-skill labor but without labor market frictions, firms cannot shed labor as Equation 28 requires the labor market to clear, that is for both firms to jointly hire all available high-skill labor for all levels of the gap $s$. This has an important effect on results as it imposes, rather mechanically, that R&D effort from skilled labor does not change with $s$ as total hiring stays the same. Hence, allowing for unemployment is important as it permits firms to adjust labor in tandem with their incentives to innovate.[51]

[51] Naturally, any framework that is isomorphic to having unemployment would also lead to similar results. I provide further details on this in Section A.6 in the Appendix.

The model also links rising skill concentration with a decline in active R&D catch-up. By making R&D imitation partly depend on high-skill labor, we can assess the link between increased skill concentration and changes to innovation diffusion. I show this in Figure A28 in the Appendix. We see that active catch-up declines as skills get concentrated at the leading firm. Importantly, this result shows how skill concentration can lead to further disincentives for the laggard: the likelihood of catching up is not only inherently small due to a high gap $s$, but also it affects the knowledge diffusion from the technological frontier to the lagging firm. In other words, it is hard to catch up because the laggard does not have enough high-skill labor. This relates to the observation in Akcigit and Ates (2023) that the growth slowdown in the US is associated with lower knowledge diffusion. Here, this happens tangibly through skilled labor.

I also assess the robustness of results to different model specifications. First, I show in Figure A29 in the Appendix the growth rate and the skill premium as a function of the gap $s$ in a version of the model where I remove the non-innovative, outside firm. Although results are slightly different from those in Figure 5, curve shapes are similar. This highlights the fact that results in the model are not being driven by the inclusion of the outside firm as most of them depend on the strategic interaction between the innovative firms. Its inclusion, however, is important for the quantitative fit of the model, particularly with respect to moments related to the labor market. As shown in Table A.22 in the Appendix, the model without the non-innovative sector struggles to match the empirical labor market tightness, cost-per-hire, and unemployment, which also affects the other moment fits. I also show robustness of results to different values of the R&D production function labor elasticity $\alpha$ in Figure A30 in the Appendix. Finally, results are robust to changes in the convexity of the R&D cost function, assumed to be quadratic in the baseline estimation. I show this in Figure A31 in the Appendix.[52]

Results remain unchanged if we change how we calculate high-skill labor concentration. As aforementioned, calculating concentration in a duopoly so as to match municipality-level data is not straightforward. A different way of doing it from my baseline method is to not use the high-skill labor employed at the non-innovative sector in the calculation of high-skill concentration. I show in Figure A33 in the Appendix that estimation results using this measure of concentration remain largely the same for both growth and the skill premium. A second alternative method is to forego the adjustment using $LC_2$, i.e. letting high-skill concentration start at 50%. I show new estimation results for this method in

[52] When estimating the model, to avoid corner solutions for some values of parameters and $s$, I add 0.005 to aggregate demand. I show in Figure A32 that results are robust to using 0.003 instead, showing that this numerical adjustment is largely innocuous.

Figure A34. As expected, concentration levels at low values of the gap $s$ are excessively high. Nonetheless, we get the same result as the baseline approach: both growth and the skill premium show a non-monotonic, inverted-U pattern as high-skill concentration increases.

As such, my model is able to capture the empirical results on non-monotonicity that we see in the data while extending the results from step-by-step innovation models to high-skill concentration at large firms. We can now use to model to assess two points. First, whether the model is able to capture the results in Section 3.1 linking an increase in high-skill supply to higher high-skill concentration, with different results on growth depending on whether high-skill concentration was high or low to begin with. Second, whether there is a role for a benevolent social planner to boost economic growth given the non-monotonicity with respect to growth. I analyze both points in the next section.

# 5 Counterfactuals

We will now use the model from Section 4 to analyze three counterfactual scenarios. In the first one, we consider the effect on growth from an increase in the aggregate supply of high-skill labor, showing that it can lead to lower growth. In the second and third ones, we analyze how Brazil could have improved its growth rate from the additional high-skill supply by propping up innovation catch-up or through a labor subsidy.

## 5.1 Counterfactual 1: Increase in High-Skill Labor Supply

As discussed in Section 1, one would usually expect the correlation between high-skill labor supply and economic growth to be positive. This is not only a consensus in public policy but it is also the expected result in several endogenous growth models. Taking Romer-based models as an example (Romer, 1990), growth increases linearly with a higher supply of human capital used in innovation (c.f. Section A.2 in the Appendix for a derivation). In the case of Brazil, high-skill supply has soared: the population share of those above 25 years old who hold a college degree has soared continuously from 5.75% in 1991 to 16.8% in 2019 (UNDP, IPEA and FJP, 2024). However, the increase in the supply of skilled labor did not seem to have boosted GDP per capita growth in the period, an unexpected result. Remarkably, college course quality did not change significantly in the period, as shown in Figure A35, nor did the composition of students among subject areas, as shown in Figure A36.[53] Although the standard Romer model does not take into account all possible mechanisms that affect the growth rate, from the model's perspective a three-fold increase in the high-skill supply should be reflected in long term growth. We can see in Figure A45 that a Romer-based model would have predicted a 63% increase in the potential growth rate between 1997 and 2019 from the increase in the high-skill share, whereas the actual growth rate trend declines.

We can, then, ask what happens to the growth rate in our model when we increase the supply of high-skill labor and whether it matches our evidence on the effect of college creation in Section 3.1. First, it is important to clarify how total labor supply (i.e. high and low skill) is affected as it not only changes aggregate demand but in reality the increase in high-skill labor is due to education which effectively converts low skill workers into high skill ones, possibly making it harder to hire the former. I do so via two scenarios. In one, labelled "external supply," I make no changes to low skill hiring and high-skill supply grows regardless (e.g. from outside sources of labor). In the other one, labelled "internal supply," low-skill hiring becomes more expense with the increase in high-skill supply as part of the skill supply comes internally due to education, i.e. that there is an additional, non-wage cost to hiring low-skill labor.[54] Second, to make comparisons easier with the data I re-estimate the model matching empirical moments for the 1999 to 2004 period except for high-skill concentration where I target the in-sample average in 2000.[55]

I then show results for the growth rate in Figure 6 along with the evolution of high-skill concentration. As we can see, the growth rate is not linearly increasing in the supply of high-skill labor as growth becomes flat at a sufficiently high $L_H$. The growth rate even declines slightly in the internal-supply case, a surprising result since this case leads to a 1:1 increase in total population as skill supply comes from abroad. This would normally result in higher growth mechanically. To compare this result with the empirical evidence, we can remove from the growth curve the population growth due to high-skill supply growth, resulting in per-capita values.[56] The per-capita growth curve has an inverted-U shape as the additional high-skill supply is not being put to use in R&D. The change in

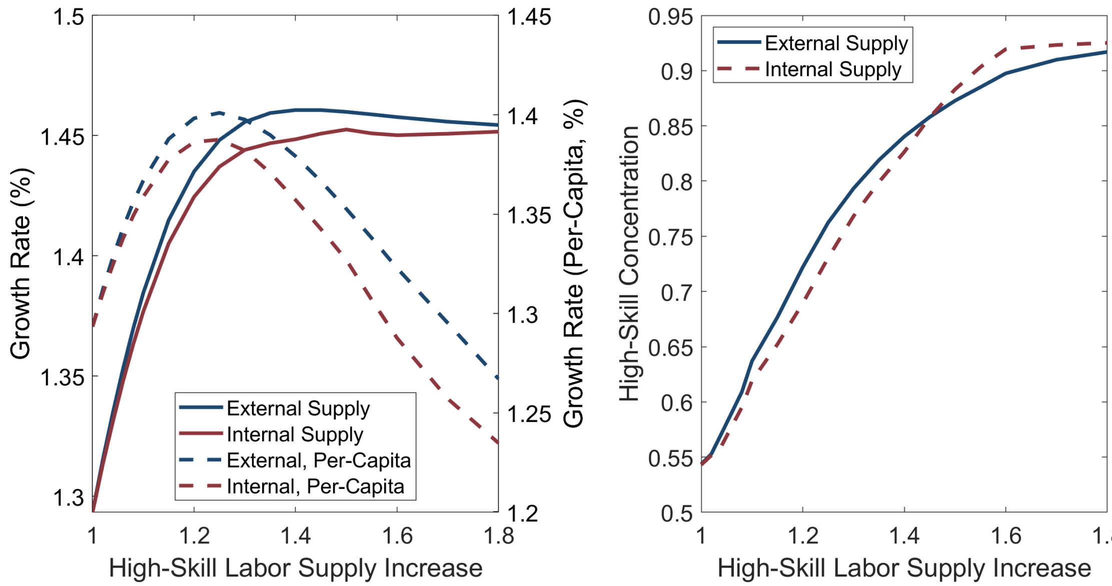
[53]The lack of change in aggregate quality might come as a surprise as we would expect *ex-ante* that the quality of the marginal student entering college to decrease with an increase in supply. A few of the reasons to why that is not the case include students facing financial constraints and having a strong distaste for distance, both of which are not necessarily correlated with student talent, and an increase in competition leading to an improvement in college quality (Cordeiro and Cox, 2023).

[54]For the realistic case, I assume hiring low-skill labor becomes more expensive as high-skill supply increases by a factor of $1 + \frac{L_H - 1}{8}$, which is increasing in $L_H$.

[55]For the moments where I do not have data in the 1999-2004 period, I use the same values as those for the whole sample, i.e. 1999-2017. I show model fit results in Table A.23 in the Appendix.

[56]For the per-capita calculation, I adjust the share of high-skill people using the empirical high-skill population share.

Figure 6: *Effect of increasing high-skill labor supply on growth and high-skill concentration for different increases in aggregate labor supply*




the growth rate slope is due to the increase in high-skill concentration and the average gap $s$ as most of the economy is now in the region where the leader innovates less as the competitive threat is diminished ("lazy monopolist"). I show this in Figure A37 in the Appendix where I compare the cross-sectional growth rates and the distribution of gaps $s$ for $L_H = 1$ (baseline) and $L_H = 1.5$. The gap distribution shifts to the right with the increase in high-skill supply as the leading firm benefits more from the decline in labor market tightness from the labor supply increase.

There are two reasons why the leader hires more high-skill labor when supply increases. The first and more direct one is due to the constant catch-up term in the follower's R&D production function. As this term is independent of high-skill labor supply, the relative increase in the follower's R&D effort is lower than the leader's. I show this in Figure A38 in the Appendix (left-hand side plot). If we assess the follower's R&D effort without the catch-up terms ("ex-h"), the lagging firm actually increases R&D effort by more than the leader at low gap levels. However, this rationale does not fully explain the increase in high-skill concentration. Even if we counterfactually increase the follower's total innovation output to match the growth of the non-catch-up part of its innovation production function, aggregate high-skill concentration still rises to 80.5% (from 53.5%).[57] We, then, still need to understand why the leader's incentive to innovate increases by more than the follower's overall.

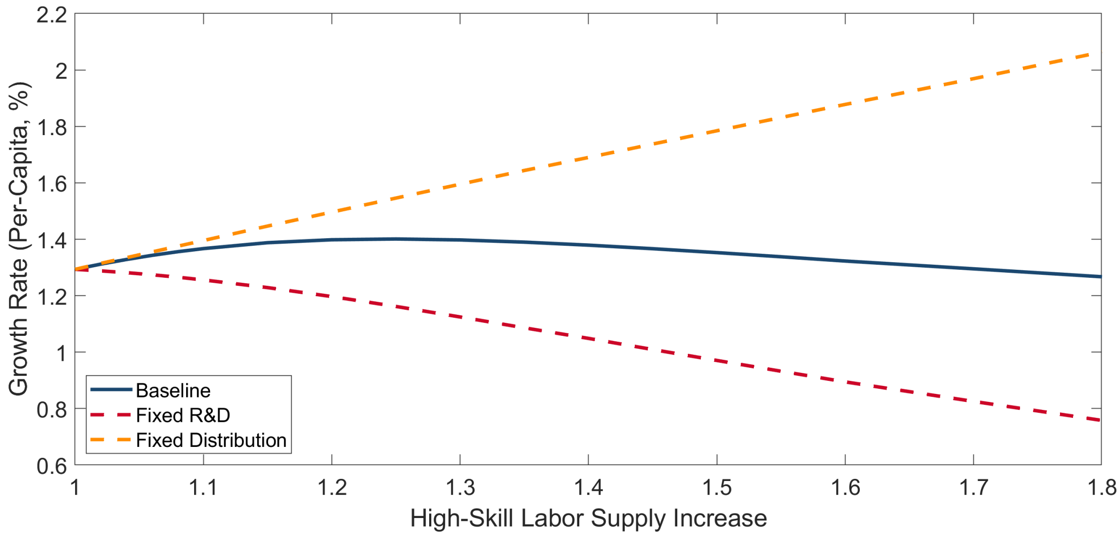
[57] To make the comparison as favorable as possible to the follower, I attribute zero growth to its R&D output where the non-catch-up term declines.

To do so, we can split up the firm's value function into two parts. In the first one, the leader derives higher value moving from $s$ to $s+1$ from being able to charge a higher markup due to the technological innovation. As the follower does not produce, this "profit-only" value flow does not matter to it. In the second one, the leader derives value from being in a better defensive position as it now takes one additional step, relative to before, for the follower to surpass the leading firm. This split applies analogously to the follower with the important remark that, for it, only this dynamic part of the value function matters. Importantly, both parts ("profit-only" and "dynamic") represent the incentive a firm has to go from $s$ to $s+1$ (or from $s$ to $s-1$, in the case of the follower).

We can then analyze what happens these two parts once high-skill supply increases. As high-skill labor becomes easier to find, hiring costs as a share of profits decline. Moreover, there is an increase in aggregate demand as an indirect effect of the increased hiring, increasing profits. This relative rise in profits, however, is lower than the relative increase in total R&D hiring. This is because the increase in profits also depends on both low-skill labor demand in production and high-skill demand at the non-innovative sector, which are only affected indirectly by the increase in high-skill supply. This implies that at a low gap level where competition is intense, the dynamic part of the value function increases by more than the profit-only part. Hence, the follower's incentive to catch-up rises by more than the leader's as the former only depends on the dynamic part. Since the follower's effort rises more in relative terms, the leader's position is at a higher threat which implies the dynamic part of its value function declines. We can see this in Figure A38 in the Appendix (right-hand plot, "L, Dynamic" vs. "F, Total") for a low gap level. The change in total incentives for the follower is larger than that for the leader at low $s$.

However, results invert at higher gap levels. As shown in Figure A38, the change in the dynamic part of the leader's value function surges at the point where the leader has the largest incentive to escape the follower's competition, i.e. the frontier between the escape-competition and the lazy-monopolist regions. This is intuitive: as the follower becomes more competitive at low $s$, the leader wants to avoid reductions of the gap more intensively. As the gap increases from that point onwards, both firms see a reduction in the dynamic incentives: for the follower, catching up becomes harder as the leader's incentives to escape competition increase, while this induces in the leader a lazy-monopolist effect due to a lower competitive threat. However, the profit-only incentive, which is exclusive of the leading firm, remains as the leader still benefits from improvements in its marginal cost. I show the breakdown between the profit-only and the dynamic parts to the total change in the leader's incentive to innovate in Figure A39 in the Appendix. We can see that the profit-only component explains why the leader increases R&D out-

Figure 7: *Decomposition of the effect of an increase in skill supply on growth*




put at higher gap levels. As such, the change in the leader's incentives from an increase in high-skill supply takes longer to decline, further lowering the follower's incentives to catch-up. This explains the rise in high-skill concentration.[58]

Though initially surprising, the decline in growth from an increase in human capital when skill concentration is high can be understood as the net effect of two channels. The first one is the boost to R&D effort when we lower the cost of innovation via a higher supply of skills, which implies higher growth. This is the usual relationship between human capital and growth in the literature. The second one is the effect on growth from shifting the gap distribution to the right (Figure A37), i.e. the overall increase in the distance between the two firms. I show both channels in Figure 7 by either fixing the initial distribution of gaps $s$ at $L_H = 1$ and allowing R&D to adjust with a larger skill supply ("Fixed R&D" curve), or by fixing the initial R&D effort and allowing the gap distribution to vary ("Fixed Distribution" curve). The total effect on growth ("Baseline") is simply the net effect of the contribution of each individual channel. Hence, this decomposition exercise makes it clear that the model does account for the usual positive effect of human capital and growth. However, it also shows how a high level of skill concentration can lead to a stronger skill concentration channel, which can more than offset the positive effect from lower R&D costs.

We can, then, assess how our model compares with the empirical results in Section 3.1. Recall that the increase in high-skill supply from college creation led to a relative decline

[58] A short way to see this is to notice that at a large enough gap, both firms have essentially no strategic incentive to innovate, yet the leading firm can still make small profit gains from improvements in marginal cost. This implies a skill concentration of 1.

in growth at highly concentrated municipalities of 10% due to a positive, short-term boost to growth where skill concentration was low and a negative, long-term decline where skill concentration was high. While the model did not target those results, it is in good measure to compare the theoretical results with our reduced-form estimates. First, it is important to highlight that the model-generated growth curve in Figure 6 is conditional on the initial level of high-skill concentration. This is intuitive: at low levels of skill concentration, more human capital boosts economic growth. At high levels, however, we observe a decline. As such, had high-skill concentration in Brazil been lower (higher), the positive-slope (negative-slope) part of the growth curve would have been longer. I show this in Figure A40 in the Appendix. Although skill concentration is an endogenous variable in the model, we can both increase and decrease its value by changing the constant catch-up parameter $h_c$.

We can now use Figure A40 to assess whether the model can capture the difference-in-differences estimates. We proceed in the following way. First, I set the initial levels of high-skill concentration for the "Low" and "High" scenarios to match the in-sample averages for the "Low Concentration" and "High Concentration" groups defined in Section 3.1, respectively. Second, we know from Figure A5 that a new college increases local skill supply, on average, by around 2 percentage-points. Relative to pre-treatment averages, this increase in high-skill supply represents, approximately, a 2.11- and a 1.77-times increase in local skill supply for the "Low" and "High" subsamples, respectively. We can then move along the growth curves to understand the reduced-form results. For the "Low Concentration" case, growth initially rises which is captured in significant and positive coefficients in Figure 2 though model results are higher than the ones estimated in reduced-form.[59] As skill concentration keeps on increasing, however, growth declines. For a 2.11-times increase in supply, the corresponding local growth rate is around 7% higher than the initial growth rate, a change that produces non-significant reduced-form coefficient estimates. As for the "High Concentration" group, growth starts to decline at a significantly lower level of high-skill supply. For most of the curve, however, growth remains nearly flat, capturing the non-significant results for the municipalities with higher skill concentration in Figure 2. At a 1.77-times increase, results imply a decline of around 7.4% relative to initial conditions, matching the long-term decline shown in my reduced-form results. Finally, the model-generated relative difference in growth between "High" and "Low" is a 14.4% decline, a reasonably close value to the reduced-form estimate of around a 10% decline. Overall, results are reassuring as the model is able to broadly

[59]The difference in levels for the short-term estimates could be due to differences between steady-state and transition dynamics as the model assumes the former while the reduced-form results capture the latter.

capture the untargeted reduced-form results in Section 3.1.[60]

Figure A40 also makes it explicit that the effect of human capital on growth crucially depends on the level of skill concentration. On one hand, in comparing the "Low" with the "Baseline" case we observe that an increase in human capital supply can have the expected positive effect on economic growth for a larger increase in skill supply if skill concentration is low. On the other hand, for a high enough level of skill concentration ("'Very High" case), per-capita growth is a monotonically decreasing function of human capital supply. We can conclude that the increase in skill supply in Brazil had a negative impact on long-term economic growth due to a combination of two things: the magnitude of the increase in supply and the initial level of skill concentration. As a corollary, it is clear that targeting high-skill concentration becomes an important policy lever to boost growth, a point which I assess in Section 5.2.

The high-skill supply increase also produces other effects in the model that we observe in the data. Regarding the skill premium and high-skill unemployment, there is a remarkable difference between the external-supply scenario and the more realistic case where hiring low-skill workers gets harder. In the former, the skill premium goes up while unemployment declines. This is due to the boost to aggregate demand from the rise in population: as high-skill supply grows, aggregate demand increases which raises profits. This, in turn, raises low-skill hiring, which further increases aggregate demand, high-skill hiring, and high-skill wages. In the latter, however, we see that results are being driven by the external-supply assumption. I show these results in Figure A41 in the Appendix. Taking the more realistic case as the benchmark, skill premium declines while high-skill unemployment go up. Both are linked to the decline in incentives to innovate from the increase in high-skill concentration. Importantly, the leading firm does not absorb the increase in $L_H$ in its entirety. Results on the skill premium and unemployment are reflected in the data, as shown in Figures A26 and A42. Results are also in line with the increase in high-skill underemployment shown in Figure A42 though the model only captures unemployment.

Finally, the model also links the increase in human capital to lower innovation diffusion. As skill supply pushes high-skill concentration up, the lagging firm engages less in active R&D imitation. I show this in Figure A43 in the Appendix for both the external and internal-supply cases. Surprisingly, even though there are more skills available in the economy, aggregate catching up by the laggard declines. This point will be a key

[60] The model also does a good job capturing the aggregate effect, i.e. for all municipalities. On average, new colleges lead to a 1.9-times increase in the local share of high-skill workers, which corresponds in the model to a 4.85% decline in growth relative to $L_H = 1$ (vs. around a 6% decline, as shown in Figure A6).

driver in Section 5.2 when we assess the role of a social planner in boosting growth. In particular, Figure A43 shows that improving knowledge diffusion from the R&D frontier to followers, either directly or indirectly via high-skill labor, is an effective measure to increase economic growth.

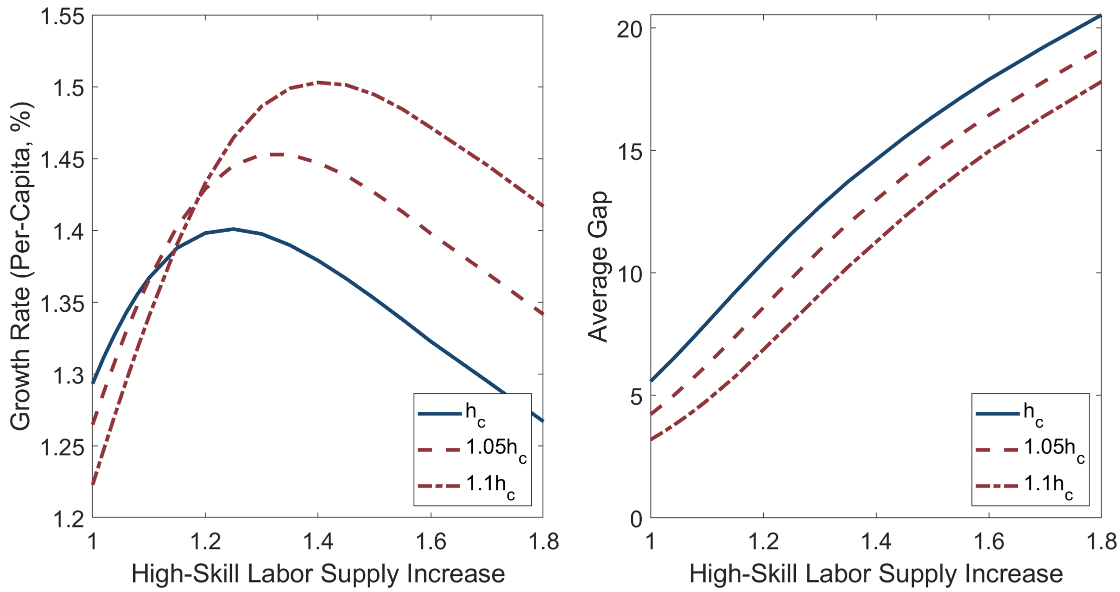
This analysis shows how we can achieve a non-linear and even a non-monotonic relationship between the growth rate and high-skill labor supply, driven by an increase in high-skill concentration. As leading firms benefit more from the increase in supply, they increase their gap relative to followers. Once the gap is high enough, incentives to innovate decline which offsets the boost to growth from a larger high-skill supply and leads to an oversupply of high-skill workers. Results are in line with both aggregate-level data in Brazil and reduced-form estimates in Section 3 on the link between high-skill concentration, growth, and the skill premium. I propose two ways a social planner could counteract the negative effect of the skill concentration channel on growth in Section 5.2.

## 5.2 Counterfactuals 2 and 3: Social Planner

After analyzing the effects of an increase in high-skill supply on growth, we can ask how a benevolent social planner could do better. A shortage of high-skill labor has been deemed one of the main obstacles for long-term growth in Brazil, prompting a government-induced increase in supply. What I showed, both through reduced-form evidence and the model, is that a larger supply of high-skill workers does not necessarily lead to more economic growth as labor concentration intensifies. However, I show next that by either targeting innovation catch-up or by subsidizing labor at the follower firm a planner is able to increase the growth rate by lowering high-skill labor concentration, weakening the negative effect of the skill concentration channel.

We start with an increase in innovation catch-up. Figure 8 shows how the response of per-capita growth rates to increases in high-skill labor supply changes if the social planner is able to increase the passive catch-up term $h_c$ from its baseline level. While $h_c$ shows up straightforwardly in the expression for the growth rate in Equation 31, the effect of increasing it on growth is ambiguous as we have to consider the interaction with high-skill concentration. By bringing the latter down, increasing the catch-up term can either increase or decrease aggregate growth rate as sectors move along the growth peak shown in Figure A37. Since the factor dampening the growth rate increase at high $L_H$ is the "lazy-monopolist" effect due to large R&D gaps, the planner can boost the growth rate by helping the follower firm catch up. This results in an increase in the growth rate at high labor supply from around 1.27% to 1.42% when $h_c$ increases by 10% as it now takes a larger increase in high-skill supply to trigger the lazy-monopolist stage.

Figure 8: *Effect of increasing high-skill labor supply for different levels of $h_c$*




In practice, we can interpret a change in the passive catch up term $h_c$ in a few different ways. First, the planner can redesign the patenting system so as to encourage disclosure of any new technologies and innovations, and to discourage firms from using (predatory) litigation as a tool to stem competition. Second, the planner can create incentives to help lagging firms catch-up to the frontier by importing intermediate goods with better quality. There is evidence that importing goods increases R&D intensity through knowledge spillovers (Chen, Zhang and Zheng, 2017). Finally, the planner could strive to lower informational barriers on best-practice "managerial technologies" (Bloom, Eifert, Mahajan, McKenzie and Roberts, 2012). As results show, there is a lot to be gained from boosting innovation catch-up when a significant amount of resources has already been spent in educating the workforce.

A second, more direct approach involves subsidizing high-skill labor at lagging firms. As high-skill supply and concentration increase, and sectors move to the lazy-monopolist state, the lagging firm stops to actively engage in R&D. That is, it stops hiring high-skill workers and investing as the odds of catching up are small and incentives to innovate are low. We can then consider the scenario where the social planner provides the follower with innovation inputs. With labor as an input to innovation, the planner can tax the leading firm and directly sponsor high-skill workers at lagging firms. Specifically, we make the following adjustments to the firms' value functions:

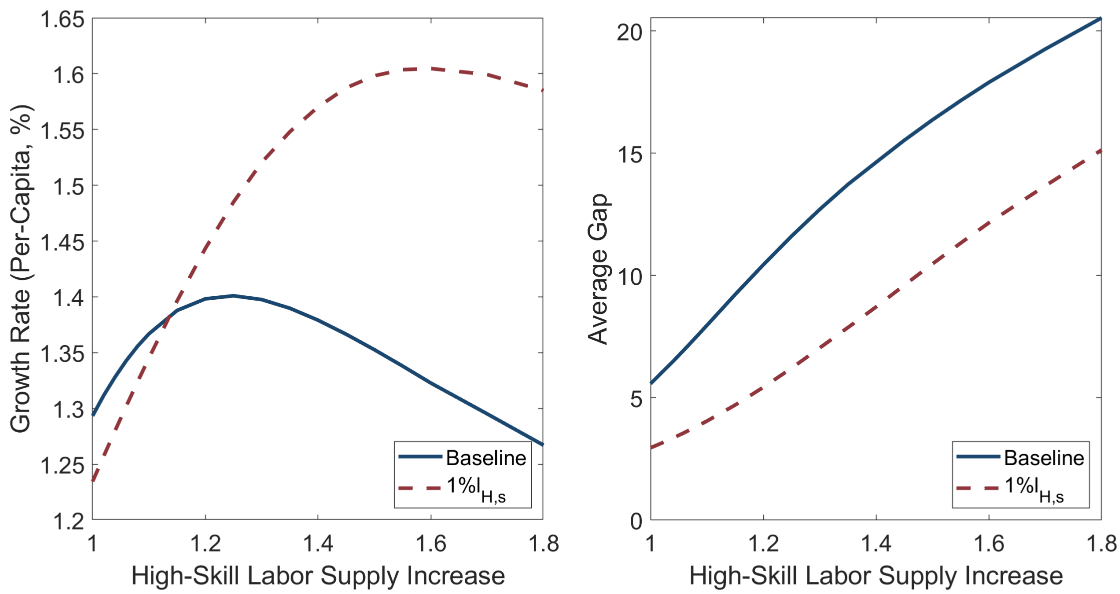
$$
\begin{aligned}
rJ_s = {} & \max_{\lambda_s, l_{s,H}} \pi_s - \rho\frac{\lambda_s^2}{2} - w_{s,H}l_{s,H}(1+\tau) - \kappa\frac{v_s^2}{2} + [A_\lambda\lambda_{-s} + A_l(l_{-s,H} + l_{s,H}\tau)^\alpha + \\
& h_l(l_{-s,H} + l_{s,H}\tau)^\alpha + h_c](J_{s-1} - J_s) + [A_\lambda\lambda_s + A_l l_{s,H}^\alpha](J_{s+1} - J_s) \\
rJ_{-s} = {} & \max_{\lambda_{-s}, l_{-s,H}} \pi_{-s} - \rho\frac{\lambda_{-s}^2}{2} - w_{-s,H}l_{-s,H} - \kappa\frac{v_{-s}^2}{2} + [A_\lambda\lambda_s + A_l l_{s,H}^\alpha](J_{-s-1} - J_{-s}) \\
& + [A_\lambda\lambda_{-s} + A_l(l_{-s,H} + l_{s,H}\tau)^\alpha + h_l(l_{-s,H} + l_{s,H}\tau)^\alpha + h_c](J_{-s+1} - J_{-s})
\end{aligned}
\tag{33}
$$

where $\tau \in [0,1]$ is the tax on the leader's total high-skill labor costs, which are then used to finance $\tau l_{s,H}$ workers at the follower.

I plot model results using the baseline estimation and $\tau = 1\%$ in Figure 9. We see that a labor subsidy is quite effective in boosting growth at large increases in high-skill supply: at an 80% increase in the latter, growth goes from around 1.45% to 1.8%. Note how the subsidized curve keeps a positive slope for longer, highlighting how the planner can recover the positive relationship between human capital and growth. This is because the subsidy helps the follower "fight back" which lower the average gap between itself and the leading firm. I show this in Figure A44 in the Appendix where I plot the firms' R&D effort for both the baseline and the subsidy cases when aggregate high-skill supply is 1.5. Although individual sectors are stochastically changing their gap $s$, the points marked with an asterisk determine points of convergence in the distribution. That is, at gap levels below the intersection the leader innovates relatively more, pushing $s$ up. However, at levels above the intersection the follower innovates relatively more, bringing $s$ back down. With the subsidy, this point of intersection moves left (from 1 to 2) to a lower gap level, indicating a lower level of high-skill concentration. Total R&D effort, and hence growth, goes up as more intense catch-up increases incentives for the leader to keep innovating in order to escape competition. This is the case even though both the leader's R&D output is lower at low levels of the gap $s$ due to the tax disincentive.

Although the increase in the growth rate from a 1% tax rate shown in Figure 9 looks impressive, it is important to understand the context here. First, the change relative to the baseline case depends on $s$ as it can turn negative for low increases in high-skill supply. This highlights that the increase in the growth rate when the skill supply increase is large comes from pent-up high-skill supply. In other words, the approximately 28% increase in growth (vs. a 2% decline in the baseline) when $L_H = 1.8$ relative to the case where $L_H = 1$ is due to the 80% increase in high-skill supply that is more appropriately being employed in innovation. Second, the tax is applied to total high-skill labor costs of the leader and it

Figure 9: *Effect of subsidizing high-skill labor at the laggard firm*




is most effective when high-skill concentration is quite high. To assess the magnitude of this tax increase, we can conduct a back-of-the-envelope calculation using US data on tax revenues and public R&D subsidies.[61] In 2019 the US government spent around $175.5 billion on R&D tax incentives and government-financed innovation. Assuming that the top 25% of the income distribution is representative of high-skill workers, an extra 1% increase in income tax amounts to around $82 billion, or almost half of all direct and indirect federal spending in R&D. The equivalent share calculated for Brazil would likely be higher as the Brazilian government spends less in R&D relative to GDP (0.82 for the US in 2019 vs. 0.4 in Brazil). Hence, the increase in R&D support would be substantial, though results show that a labor subsidy to innovation in lagging firms that takes into account high-skill concentration can be quite effective in boosting growth.

Importantly, this analysis points towards a different direction regarding education policy in places where skill concentration is high. An ever-increasing high-skill labor supply, in itself, is not a recipe for higher growth rates once the skill concentration channel dominates the positive effect of human capital on growth. What is key to this conclusion is understanding the interaction between the high-skill labor market and how innovative firms compete in the R&D space. As such, calls for a higher supply of skills should be understood within the context of high-skill labor concentration at large firms. Along with boosting skill supply, government should also focus on competition policy, particularly

[61] Along with being an easier reference to most people, the US tax data is more easily available. Data comes from the Tax Foundation and the OECD for public R&D spending.

within-sector innovation catch-up and diffusion.

# 6 Conclusion

I show in this paper how the effect of human capital on growth depends on high-skill labor concentration at large firms. I start by showing causal evidence in a difference-in-differences design that increases to local skill supply from college creation had a negative and significant effect on growth in municipalities where high-skill workers across firms was high. Results are robust to different specifications, changes to the sample, and show a relative decline of around 10% in local growth between places with high and low skill concentration in the long term.

I then proceed to establish the role of high-skill concentration in the link between human capital and growth. First, I leverage the same difference-in-differences approach to show that the increase in local skill supply led to an increase in local high-skill concentration by around 12%. Second, I build an SSIV using data on public loans to firms to show that local high-skill concentration is non-monotonically related to local GDP growth. At low levels of concentration, increasing high-skill labor at large firms boosts economic growth. If skill concentration keeps increasing, the relationship inverts and growth starts to decline. I further show causal evidence using the same SSIV that local skill concentration also has a non-monotonic relationship with local skill premium. I show that my identification strategy passes the recommended tests in the literature of shift-share IVs, and that estimates are robust to several changes in the specification.

I then rationalize results in a model with step-by-step innovation and high-skill labor demand and search. When firms are close in the technology ladder, competition is intense which raises the growth as firms compete for R&D labor input. Once a leader is significantly far ahead, it reduces its innovation effort as the threat of competition is lower and the likelihood of a lagging firm catching up is low. All the while, I show that high-skill concentration at the leader is monotonically increasing in the R&D gap. Thus, the model is able to reproduce the non-monotonic relationships observed in the data between skill concentration and growth.

With the model in hands, I analyze the effect of an increase in skill supply on growth. I show that this effect can be decomposed into two parts: one positive, due to the boost to R&D effort due to lower high-skill hiring costs, and one negative, due to the increase in skill concentration across firms. I further show that the negative effect more than offsets the positive one when the level of skill concentration is high enough, leading to a decline in growth. Results on growth also match the reduced-form evidence on the relative decline in growth in highly concentrated municipalities from an increase in high-skill supply. The model also captures the decline in aggregate skill premium and the rise of high-skill unemployment in Brazil.

I then assess the role of a social planner in boosting growth after increasing high-skill labor supply in places where skill concentration is high. I show that once the planner helps the lagging firm catch-up, either through a increase in technology diffusion or through a subsidy to high-skill labor at the follower, they can effectively counteract the high-skill concentration channel and increase growth. This is relevant as it highlights the important role that firm dynamics and interaction should play in education policy as increasing skill supply when skill concentration is high can effectively backfire as the policy ends up helping large firms grow even larger. As such, both education policy and competition policy should go hand-in-hand.

My results, then, show that raising high-skill supply increases high-skill labor concentration at large firms and can lead to lower growth. Moreover, results are able to explain several of the observed empirical regularities in Brazil. By focusing on high-skill labor concentration, I am able to explain the puzzling observation that a three-fold increase in high-skill labor supply did not produce an increase in growth trends in Brazil between the late 1990's and the 2010's. My model also proposes a micro-foundation to the low business dynamism observed in Akcigit and Ates (2023) for the US. As firms away from the technology frontier require high-skill labor to catch-up, an increase in labor market power at the leading firm could make it harder for a lagging firm to adopt innovation from the frontier. Crucial to this point is seeing high-skill labor flows as a channel for knowledge diffusion between firms. This is related to the use of non-compete clauses in the US where a firm can block knowledge flows by blocking former employees from being hired by competitors.

While not in the scope of this paper, I leave two ideas for future work. First, this framework can be easily expanded to take into account inter-sector labor market competition. A sector leader who experiences an increase in its labor search productivity can reduce R&D effort in other sectors competing for similar workers. It would be interesting to understand the role of high-skill worker concentration and hiring competition in explaining structural shifts in the economy, for example from manufacturing to services. Moreover, the model can also be applied in the context of competition between a domestic ("laggard") and a foreign ("leader") firm. Through "brain-drain" where domestic high-skill workers go to work at market leaders abroad, domestic firms may find themselves unable to keep up with the technological frontier. The same rationale can be applied within a country between two regions where high-skill labor migrates from one region to another.

# References


**Abramovsky, Laura, Rupert Harrison, and Helen Simpson**, "University Research and the Location of Business R&D," *The Economic Journal*, 2007, *117* (519), C114–C141.

**Acemoglu, Daron and Ufuk Akcigit**, "Intellectual Property Rights Policy, Competition and Innovation," *Journal of the European Economic Association*, 2012, *10* (1), 1–42.

**Adão, Rodrigo, Michal Kolesár, and Eduardo Morales**, "Shift-Share Designs: Theory and Inference," *The Quarterly Journal of Economics*, 2019, *134* (4), 1949–2010.

**Aghion, Philippe and Peter Howitt**, "A Model of Growth Through Creative Destruction," *Econometrica*, 1992, *60* (2), 323–351.

**_, Christopher Harris, and John Vickers**, "Competition and growth with step-by-step innovation: An example," *European Economic Review*, 1997, *41* (3-5), 771–782.

**_, _, Peter Howitt, and John Vickers**, "Competition, Imitation and Growth with Step-by-Step Innovation," *The Review of Economic Studies*, 2001, *68* (3), 467–492.

**Akcigit, Ufuk and Nathan Goldschlag**, "Where Have All the "Creative Talents" Gone? Employment Dynamics of US Inventors," Working Paper 31085, National Bureau of Economic Research 2023.

**_ and Sina T. Ates**, "What Happened to US Business Dynamism?," *Journal of Political Economy*, 2023, *131* (8), 2059–2124.

**Arrow, Kenneth J.**, "The Economic Implications of Learning by Doing," *The Review of Economic Studies*, 1962, *29* (3), 155–173.

**Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen**, "The Fall of the Labor Share and the Rise of Superstar Firms," *The Quarterly Journal of Economics*, 2020, *135* (2), 645–709.

**Azar, José, Ioana Marinescu, and Marshall Steinbaum**, "Labor Market Concentration," *Journal of Human Resources*, 2022, *57* (S), S167–S199.

**Azoulay, Pierre, Joshua S Graff Zivin, Danielle Li, and Bhaven N Sampat**, "Public R&D investments and private-sector patenting: evidence from NIH funding rules," *The Review of economic studies*, 2019, *86* (1), 117–152.

**Becker, Gary S.**, "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 1962, *70* (5), 9–49.

**Benhabib, Jess and Mark M. Spiegel**, "Human Capital and Technology Diffusion," in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Vol. 1, Elsevier, 2005, chapter 13, pp. 935–966.

**Benmelech, Efraim, Nittai K. Bergman, and Hyunseob Kim**, "Strong Employers and Weak Employees," *Journal of Human Resources*, 2022, *57* (S), S200–S250.

**Betarelli Junior, Admir Antonio, Weslem Rodrigues Faria, Rosa Livia Gonçalves Montenegro, Domitila Santos Bahia, and Eduardo Gonçalves**, "Research and development, productive structure and economic effects: Assessing the role of public financing in Brazil," *Economic Modelling*, 2020, *90*, 235–253.

**Bianchi, Nicola and Michela Giorcelli**, "Scientific Education and Innovation: From Technical Diplomas to University Stem Degrees," *Journal of the European Economic Association*, 2019, *18* (5), 2608–2646.

**Biasi, Barbara and Song Ma**, "The Education-Innovation Gap," Working Paper 29853, National Bureau of Economic Research 2022.

**Bils, Mark and Peter J Klenow**, "Does schooling cause growth?," *American economic review*, 2000, *90* (5), 1160–1183.

**Black, Sandra E. and Lisa M. Lynch**, "Human-Capital Investments and Productivity," *The American Economic Review*, 1996, *86* (2), 263–267.

**Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts**, "Does Management Matter? Evidence from India," *The Quarterly Journal of Economics*, 2012, *128* (1), 1–51.

**Borusyak, Kirill, Peter Hull, and Xavier Jaravel**, "Quasi-Experimental Shift-Share Research Designs," *The Review of Economic Studies*, 2021, *89* (1), 181–213.

**Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo**, "Capital Utilization and Returns to Scale," *NBER macroeconomics annual*, 1995, *10*, 67–110.

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

**Cattaneo, Matias D., Richard K. Crump, Max H. Farrell, and Yingjie Feng**, "On Binscatter," *American Economic Review*, May 2024, *114* (5), 1488–1514.

**Che, Yi and Lei Zhang**, "Human Capital, Technology Adoption and Firm Performance: Impacts of China's Higher Education Expansion in the Late 1990s," *The Economic Journal*, 2017, *128* (614), 2282–2320.

**Chen, Zhiyuan, Jie Zhang, and Wenping Zheng**, "Import and innovation: Evidence from Chinese firms," *European Economic Review*, 2017, *94*, 205–220.

**Cohen, Wesley M. and Daniel A. Levinthal**, "Innovation and Learning: The Two Faces of R&D," *The Economic Journal*, 1989, *99* (397), 569–596.

**Cordeiro, Fernando and Alvaro Cox**, "College Quality and Tuition Subsidies in Equilibrium," *Working Paper*, 2023.

**Cox, Alvaro**, "From Classroom to Prosperity: Fostering Development Through Higher Education," *Working Paper*, 2024.

**Damodaran, Aswath**, "Equity risk premiums (ERP): Determinants, estimation and implications - The 2023 edition," *NYU Stern School of Business*, 2023.

**de Chaisemartin, Clément and Xavier D'Haultfœuille**, "Difference-in-Differences Estimators of Intertemporal Treatment Effects," *The Review of Economics and Statistics*, 2024, pp. 1–45.

**Diamond, Peter A.**, "Wage Determination and Efficiency in Search Equilibrium," *The Review of Economic Studies*, 1982, *49* (2), 217–227.

**Dix-Carneiro, Rafael and Brian K. Kovak**, "Trade Liberalization and the Skill Premium: A Local Labor Markets Approach," *American Economic Review*, May 2015, *105* (5), 551–57.

**_ , Pinelopi K Goldberg, Costas Meghir, and Gabriel Ulyssea**, "Trade and Informality in the Presence of Labor Market Frictions and Regulations," Working Paper 28391, National Bureau of Economic Research January 2021.

**Döttling, Robin, German Gutierrez Gallardo, and Thomas Philippon**, "Is there an investment gap in advanced economies? If so, why?," *ECB Forum on Central Banking*, 2017.

**Durlauf, Steven N., Paul A. Johnson, and Jonathan R.W. Temple**, "Chapter 8 Growth Econometrics," in Philippe Aghion and Steven N. Durlauf, eds., *Philippe Aghion and*

*Steven N. Durlauf, eds.*, Vol. 1 of *Handbook of Economic Growth*, Elsevier, 2005, pp. 555–677.

**Evenson, Robert E. and Yoav Kislev**, "Research and Productivity in Wheat and Maize," *Journal of Political Economy*, 1973, *81* (6), 1309–1329.

**Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer**, "The Next Generation of the Penn World Table," *American Economic Review*, October 2015, *105* (10), 3150–82.

**Felix, Mayara**, "Trade, labor market concentration, and wages," *Job Market Paper*, 2022.

**Growiec, Jakub, Peter McAdam, and Jakub Muck**, "R&D Capital and the Idea Production Function," Research Working Paper RWP 23-05, Federal Reserve Bank of Kansas City May 2023.

**Hall, Bronwyn and Josh Lerner**, "The Financing of R&D and Innovation," in "Handbook of the Economics of Innovation," Vol. 1, Elsevier, 2010, chapter Chapter 14, pp. 609–639.

**Hansen, Lars Peter**, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 1982, *50* (4), 1029–1054.

**Hanushek, Eric A. and Dennis D. Kimko**, "Schooling, Labor-Force Quality, and the Growth of Nations," *American Economic Review*, December 2000, *90* (5), 1184–1208.

**Harrigan, James, Ariell Reshef, and Farid Toubal**, "Techies and Firm Level Productivity," Working Paper 31341, National Bureau of Economic Research 2023.

**Hausman, Jerry A., Whitney K. Newey, Tiemen Woutersen, John C. Chao, and Norman R. Swanson**, "Instrumental variable estimation with heteroskedasticity and many instruments," *Quantitative Economics*, 2012, *3* (2), 211–255.

**Hausman, Naomi**, "University Innovation and Local Economic Growth," *The Review of Economics and Statistics*, 2022, *104* (4), 718–735.

**Hopenhayn, Hugo, Julian Neira, and Rish Singhania**, "From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share," *Econometrica*, 2022, *90* (4), 1879–1914.

**Iacus, Stefano M., Gary King, and Giuseppe Porro**, "Causal Inference without Balance Checking: Coarsened Exact Matching," *Political Analysis*, 2012, *20* (1), 1–24.

**Jaeger, David A, Joakim Ruist, and Jan Stuhler**, "Shift-Share Instruments and the Impact of Immigration," Working Paper 24285, National Bureau of Economic Research 2018.

**Jarosch, Gregor, Jan Sebastian Nimczik, and Isaac Sorkin**, "Granular Search, Market Structure, and Wages," *The Review of Economic Studies*, 01 2024, p. rdae004.

**Jones, Charles I.**, "R & D-Based Models of Economic Growth," *Journal of Political Economy*, 1995, *103* (4), 759–784.

**Lameiras, Maria Andreia Parente and Leonardo Siqueira Vasconcelos**, "A evolução da população ocupada com nível superior no mercado de trabalho," Technical Report 41, IPEA 2018.

**Liu, Ernest, Atif Mian, and Amir Sufi**, "Low Interest Rates, Market Power, and Productivity Growth," *Econometrica*, 2022, *90* (1), 193–221.

**Lobel, Felipe**, "Who Benefits from Payroll Tax Cuts? Market Power, Tax Incidence and Efficiency," *Job Market Paper*, 2024.

**Loecker, Jan De, Jan Eeckhout, and Simon Mongey**, "Quantifying market power and business dynamism in the macroeconomy," Technical Report, National Bureau of Economic Research 2021.

**Lucas, Robert E.**, "On the mechanics of economic development," *Journal of Monetary Eco-*

*nomics*, 1988, *22* (1), 3–42.

**Manera, Andrea**, “Competing for Inventors: Market Concentration and the Misallocation of Innovative Talent,” *Available at SSRN 4212815*, 2022.

**Mankiw, N.G., David Romer, and David Weil**, “A Contribution to the Empirics of Economic Growth,” *Quarterly Journal of Economics*, 1992, *107* (May), 407–437.

**Moody, Carlisle E**, “The Measurement of Capital Services by Electrical Energy,” *Oxford Bulletin of Economics and Statistics*, 1974, *36* (1), 45–52.

**Moris, F and B Shackelford**, “Labor Costs Account for Over Two-Thirds of U.S. Business R&D Performance in 2020,” Technical Report 23-322, National Science Foundation 2023.

**Mortensen, Dale**, “The Matching Process as a Noncooperative Bargaining Game,” in “The Economics of Information and Uncertainty,” National Bureau of Economic Research, Inc, 1982, pp. 233–258.

**Nelson, Richard R. and Edmund S. Phelps**, “Investment in Humans, Technological Diffusion, and Economic Growth,” *The American Economic Review*, 1966, *56* (1/2), 69–75.

**Nimier-David, Elio**, “Local Human Capital and Firm Creation: Evidence from the Massification of Higher Education in France,” *Working Paper*, 2023.

**Olea, José Luis Montiel and Carolin Pflueger**, “A Robust Test for Weak Instruments,” *Journal of Business & Economic Statistics*, 2013, *31* (3), 358–369.

**Olmstead-Rumsey, Jane**, “Market Concentration and the Productivity Slowdown,” Manuscript 2022.

**Oster, Emily**, “Unobservable Selection and Coefficient Stability: Theory and Evidence,” *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

**Pessoa, Argentino**, ““Ideas” driven growth: the OECD evidence,” *Portuguese Economic Journal*, Apr 2005, *4* (1), 46–67.

**Petrongolo, Barbara and Christopher A. Pissarides**, “Looking into the Black Box: A Survey of the Matching Function,” *Journal of Economic Literature*, 2001, *39* (2), 390–431.

**Pires, Manoel, Pedro Romero Marques, and José Bergamin**, “A tributação da renda corporativa no Brasil: estimativas da carga tributária efetiva a partir das demonstrações de resultado no período 2012-2022,” Research Working Paper, FGV 2023.

**Pissarides, Christopher A.**, “Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages,” *The American Economic Review*, 1985, *75* (4), 676–690.

**Pritchett, Lant**, “Chapter 11 Does Learning to Add up Add up? The Returns to Schooling in Aggregate Data,” in E. Hanushek and F. Welch, eds., *E. Hanushek and F. Welch, eds.*, Vol. 1 of *Handbook of the Economics of Education*, Elsevier, 2006, pp. 635–695.

**Ridder, Maarten De**, “Market power and innovation in the intangible economy,” *American Economic Review*, 2024, *114* (1), 199–251.

**Romer, Paul M.**, “Endogenous Technological Change,” *Journal of Political Economy*, 1990, *98* (5), S71–S102.

**Roseth, Viviana V, Alexandria Valerio, Marcela Gutiérrez et al.**, “Education, Skills and Labor Market Outcomes,” *World Bank Group. Retrieved February*, 2016, *15*, 2017.

**Roth, Jonathan, Pedro H.C. Sant’Anna, Alyssa Bilinski, and John Poe**, “What’s trending in difference-in-differences? A synthesis of the recent econometrics literature,” *Journal of Econometrics*, 2023, *235* (2), 2218–2244.

**Schubert, Gregor, Anna Stansbury, and Bledi Taska**, “Employer concentration and outside options,” *Available at SSRN 3599454*, 2024.

**Sun, Liyang and Sarah Abraham**, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 2021, *225* (2), 175–199.

**Taylor, Jim**, "A Surrogate for Regional Estimates of Capital Stock," *Bulletin of the Oxford University Institute of Economics & Statistics*, 1967, *29* (3), 289–299.

**Tilton, John E.**, "International diffusion of technology: the case of semiconductors," 1971.

**Toivanen, Otto and Lotta Vaananen**, "Education and Invention," *The Review of Economics and Statistics*, 2016, *98* (2), 382–396.

**Ulyssea, Gabriel**, "Regulation of entry, labor market institutions and the informal sector," *Journal of Development Economics*, 2010, *91* (1), 87–99.

**UNDP, IPEA, and FJP**, "Atlas do Desenvolvimento Humano no Brasil," http://www.atlasbrasil.org.br/ 2024. Accessed June 12, 2024.

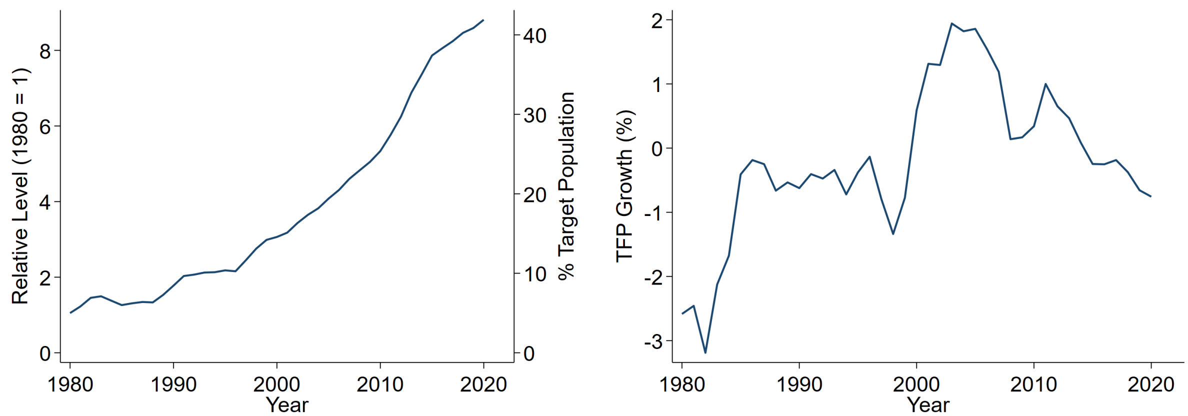
**Valero, Anna and John Van Reenen**, "The economic impact of universities: Evidence from across the globe," *Economics of Education Review*, 2019, *68*, 53–67.

**Veloso, Fernando, Fernando de Holanda Barbosa Filho, and Paulo Peruchetti**, "Impactos da educação no mercado de trabalho," Technical Report, IBRE - FGV 2022.

# A Appendix

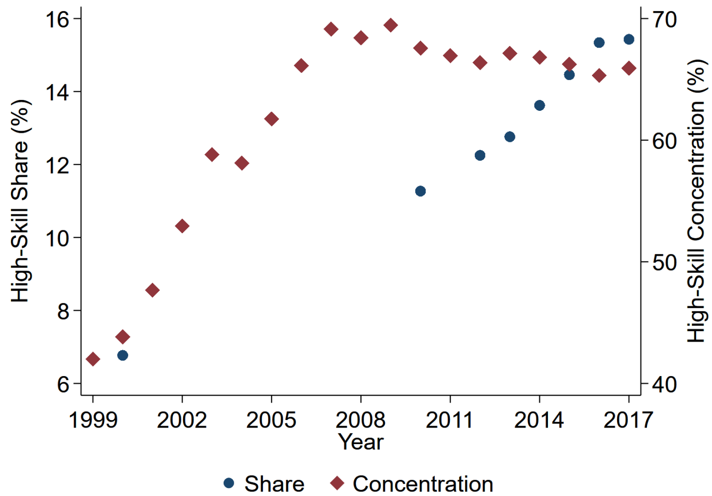
## A.1 Figures and Tables

Figure A1: *Evolution of college enrollment and TFP growth in middle-income countries*



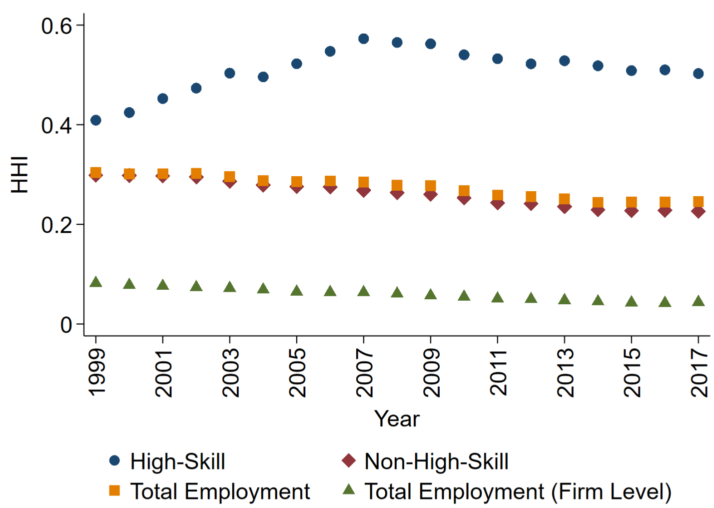

Note: College enrollment data comes from the World Bank. Target population is the population of the age group which officially corresponds to college education. Series show 3-year moving averages and are weighted by country population. TFP growth data is from the Penn World Table (Feenstra, Inklaar and Timmer, 2015).

Figure A2: *Evolution of high-skill share in the population and local skill concentration*




Note: High-skill share data is from the Atlas of Human Development in Brazil (UNDP et al., 2024). High-skill share corresponds to the ratio between the number of people with a college degree and the total population who is at least 25 years old. High-skill concentration is the median across municipalities of the local share of high-skill people working at large firms over total local supply.

Figure A3: *Evolution of the HHI-style concentration for different types of labor*




Note: High-skill workers are those with at least some college education, though they might not have finished their degree. The HHI-based measure of concentration for high-skill, non-high-skill, and total employment is calculated by splitting firms by size bins and calculating the employment HHI between those bins, i.e. by using bin employment shares. To avoid cases where bins of smaller firms have more employees than those of larger firms, I drop localities where that happens. Total Employment (Firm Level) shows an HHI measure calculated using firm-level employment shares.

Figure A4: *Evolution in the number of colleges and the share of college graduates in the population*

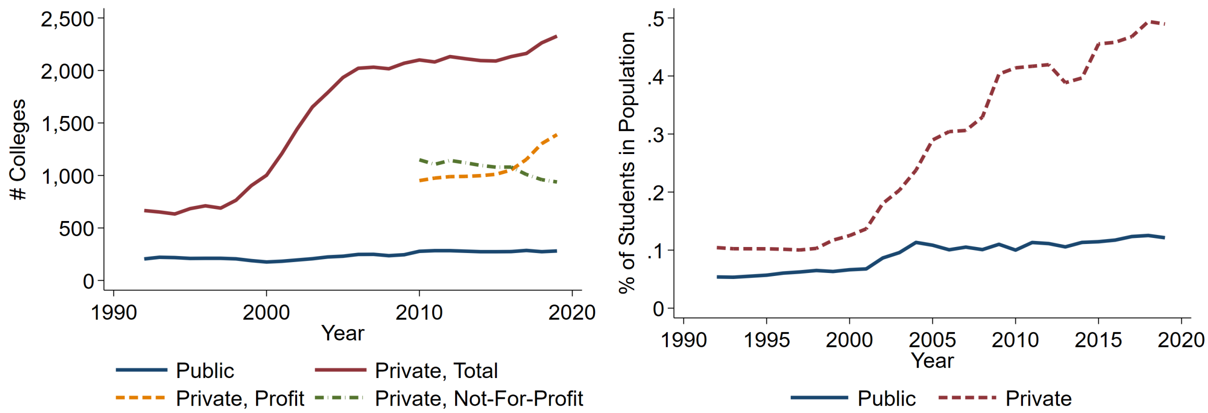
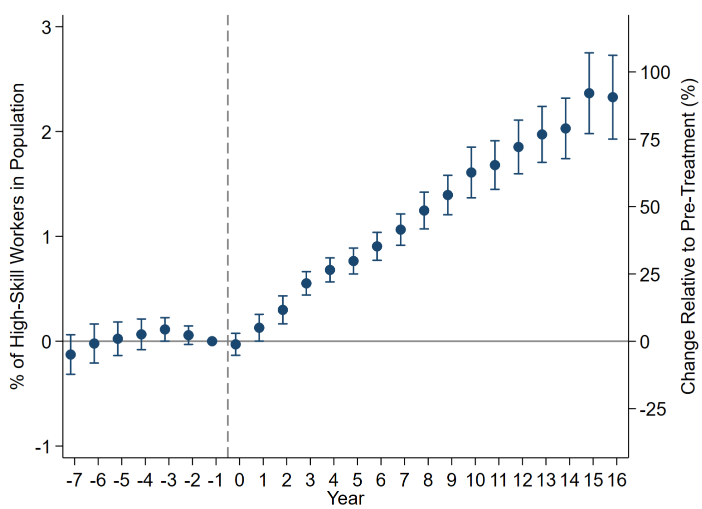
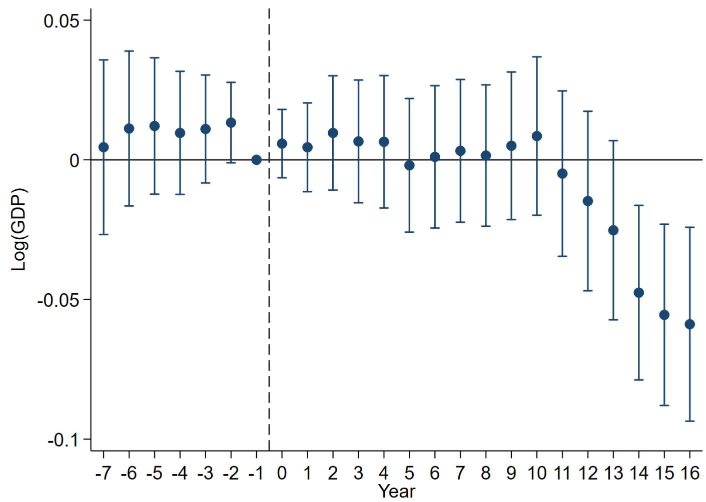



Figure A5: *Effect of college creation on local share of high-skill people*



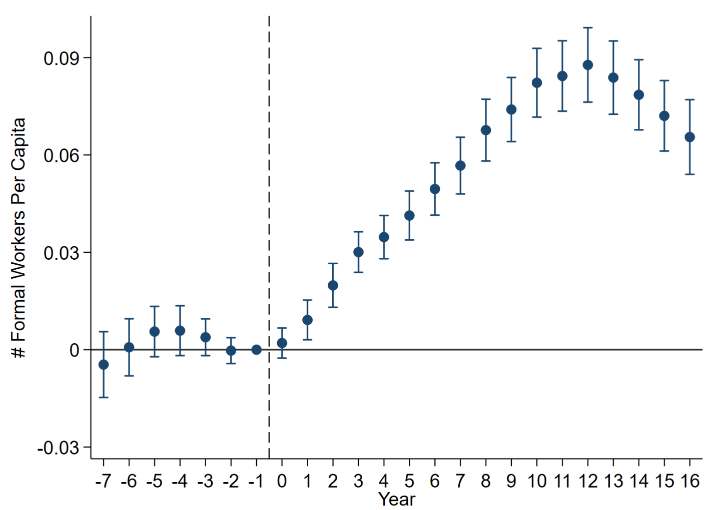

Note: Change relative to pre-treatment period shows the ratio between the coefficient estimates and the average share of high-skill people across untreated municipalities. Vertical bars represent the 95% confidence interval.

Figure A6: *Difference-in-differences estimates of the effect of college creation on local growth*



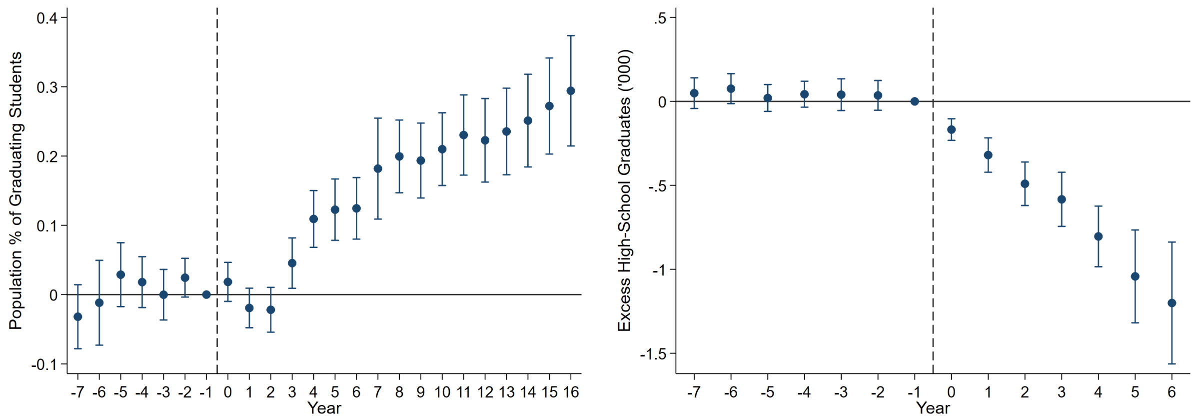

Vertical bars represent the 95% confidence interval.

Figure A7: *Difference-in-differences estimates of the effect of college creation on local formal employment*



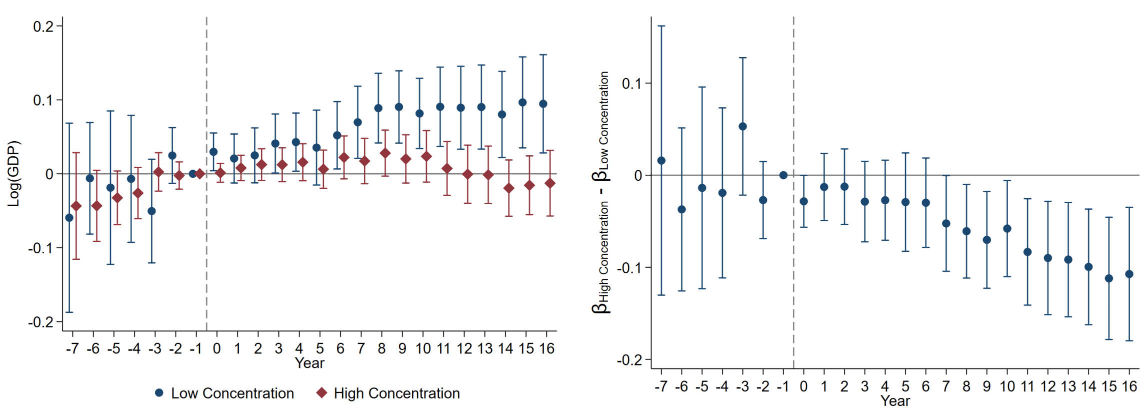

Employment data uses employer-employee links to calculate the local stock of formal workers. Vertical bars represent the 95% confidence interval.

Figure A8: *Trends on college supply competition and college demand relative to the arrival of a new college*



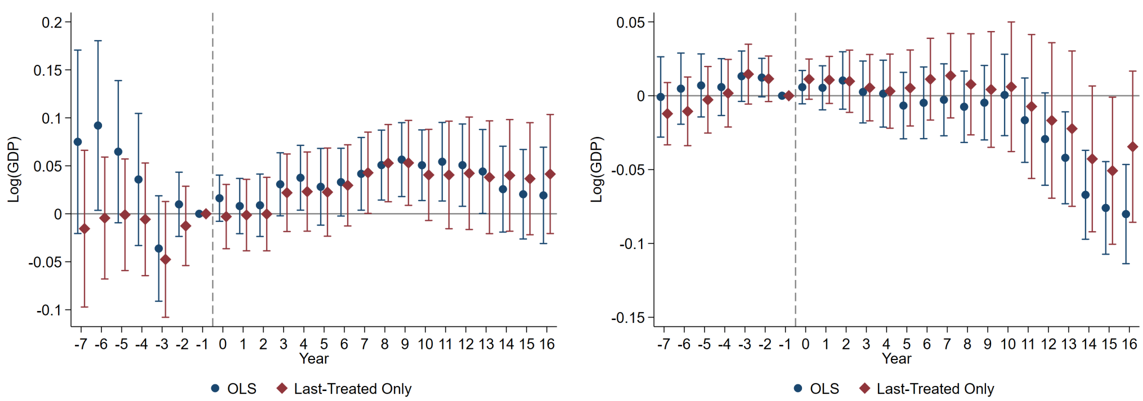

Population % of graduating students refers to the population share of college students who graduated in each year. Excess high-school graduates refers to the difference between the number of new high-school graduates in each year and the size of the incoming first-year college cohort, in thousands. High-school data comes from INEP and is restricted to the 1999-2006 period. Vertical bars represent the 95% confidence interval.

Figure A9: *Difference-in-differences estimates of the effect of college creation on local growth at municipalities with high and low high-skill concentration using the placebo-to-treated matched sample*



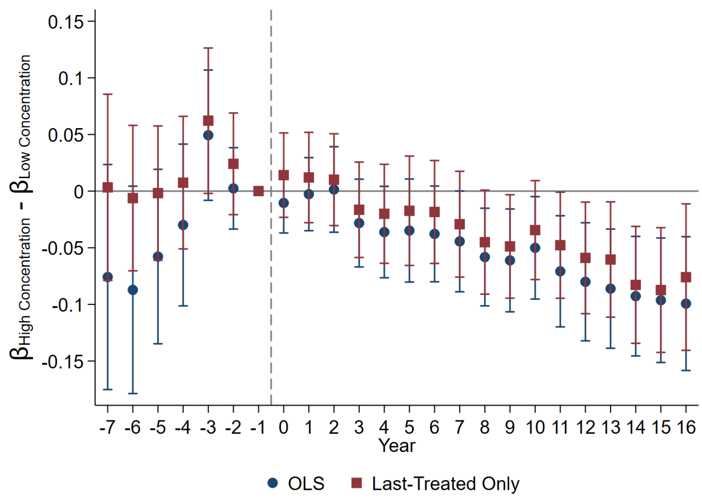

Note: High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Sample match is on population level, share earning minimum wage or lower, share who only completed the $5^{th}$ grade, unemployment rate, and illiteracy rate, all in 2000. Treated observations are matched to those in control using the coarsened matching method in Iacus, King and Porro (2012). Low (high) concentration municipalities are defined as those below (above) the $14^{th}$ percentile of high-skill concentration averaged within the three-year period pre-treatment, for those treated, or below (above) the same concentration level averaged in the initial two periods for the non-treated so as to get similar threshold levels in both cases. Vertical bars represent the 90% confidence interval.

Figure A10: *Difference-in-differences estimates of the effect of college creation on local growth at municipalities with high and low high-skill concentration using the last-treated control group*



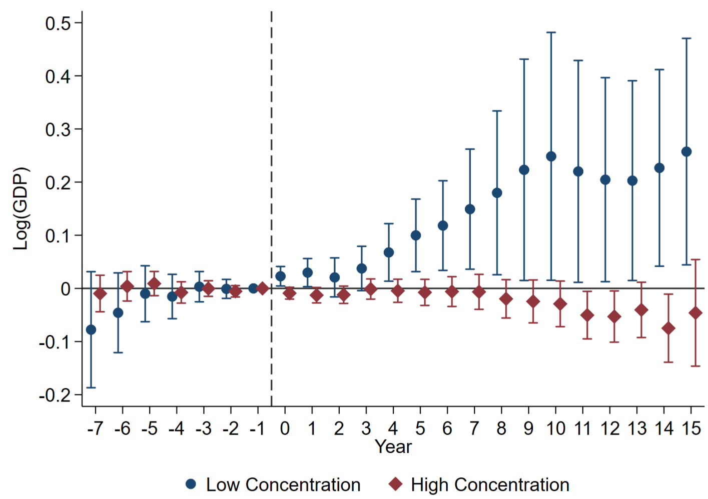

Note: High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Low (high) concentration municipalities are defined as those below (above) the $14^{th}$ percentile of high-skill concentration averaged within the three-year period pre-treatment, for those treated, or below (above) the same concentration level averaged in the initial two periods for the non-treated so as to get similar threshold levels in both cases. Estimates restrict the control group to last-treated units and use the estimator proposed in Sun and Abraham (2021). Vertical bars represent the 90% confidence interval.

Figure A11: *Difference-in-differences estimates of the difference in the effect of college creation on local growth between municipalities with high and low high-skill concentration using the last-treated control group*



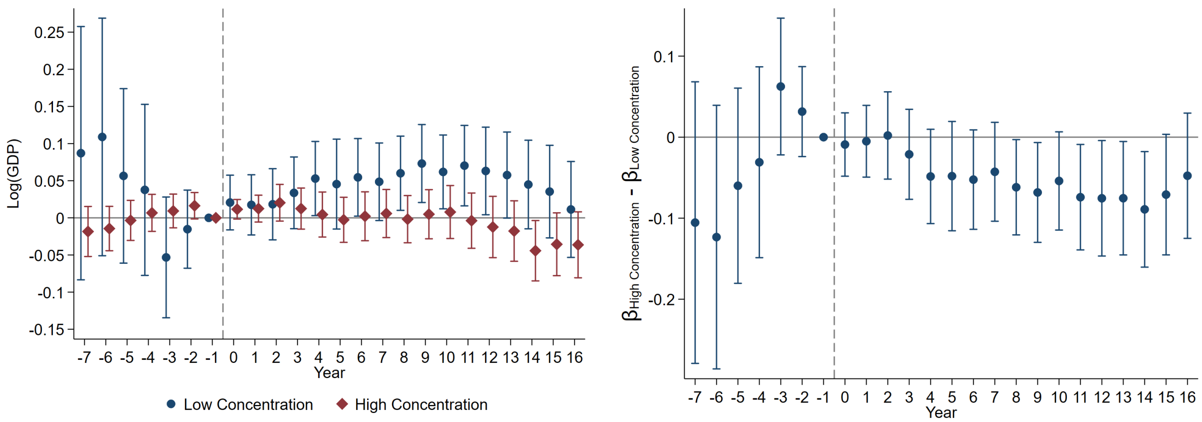

Note: High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Low (high) concentration municipalities are defined as those below (above) the $14^{th}$ percentile of high-skill concentration averaged within the three-year period pre-treatment, for those treated, or below (above) the same concentration level averaged in the initial two periods for the non-treated so as to get similar threshold levels in both cases. Estimates restrict the control group to last-treated units and use the estimator proposed in Sun and Abraham (2021). Vertical bars represent the 90% confidence interval.

Figure A12: *Difference-in-differences estimates of the effect of college creation on local growth at municipalities with high and low high-skill concentration using the not-yet-treated control group*



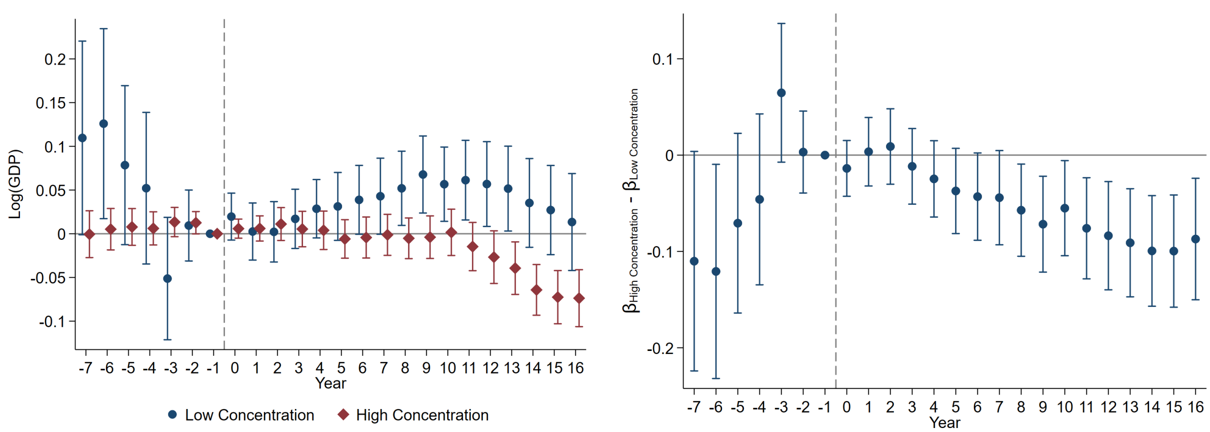

Note: High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Low (high) concentration municipalities are defined as those below (above) the $20^{th}$ percentile of high-skill concentration averaged within the three-year period pre-treatment, for those treated, or below (above) the same concentration level averaged in the initial two periods for the non-treated so as to get similar threshold levels in both cases. Estimates restrict the control group to not-yet-treated units and use the estimator proposed in de Chaisemartin and D'Haultfœuille (2024). Vertical bars represent the 90% confidence interval.

Figure A13: *Difference-in-differences estimates of the effect of college creation on local growth at municipalities with high and low high-skill concentration for the no-college sample*



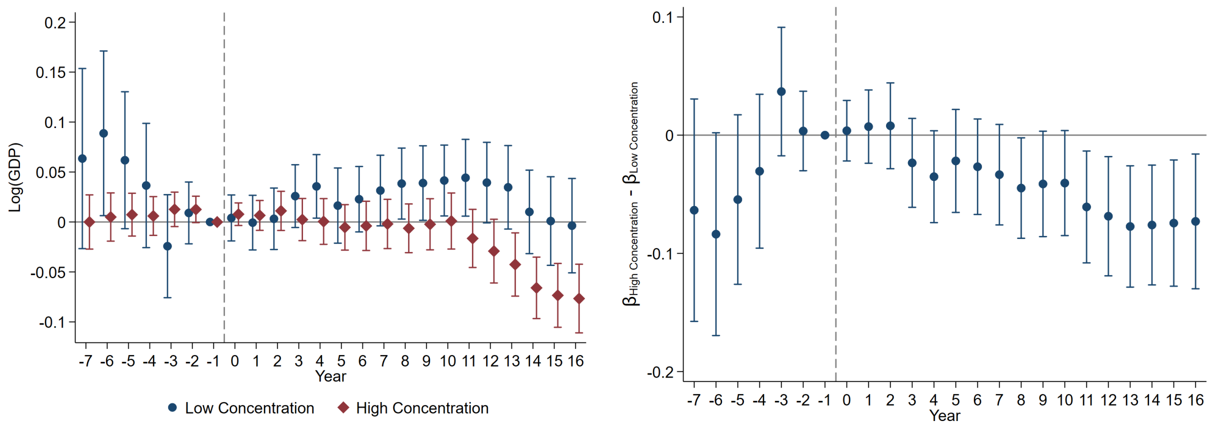

Note: No-college sample only includes observations that do not have a college in all periods (control) or in the pre-treatment period (treated). High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Low (high) concentration municipalities are defined as those below (above) the $14^{th}$ percentile of high-skill concentration averaged within the three-year period pre-treatment, for those treated, or below (above) the same concentration level averaged in the initial two periods for the non-treated so as to get similar threshold levels in both cases and between the no-college sample and the baseline one. Vertical bars represent the 90% confidence interval.

Figure A14: *Difference-in-differences estimates of the effect of college creation on local growth at municipalities with high and low high-skill concentration for* $p = 12\%$



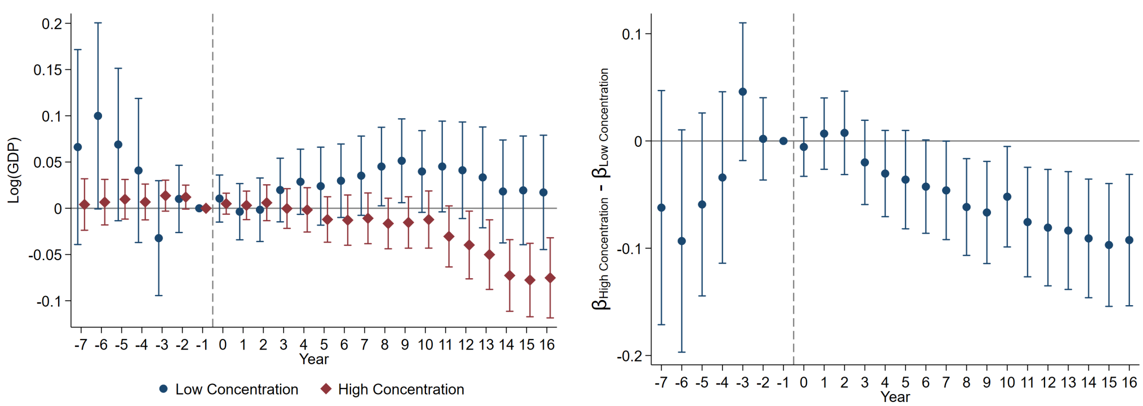

Note: High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Low (high) concentration municipalities are defined as those below (above) the $12^{th}$ percentile of high-skill concentration averaged within the three-year period pre-treatment, for those treated, or below (above) the same concentration level averaged in the initial two periods for the non-treated so as to get similar threshold levels in both cases. Vertical bars represent the 90% confidence interval.

Figure A15: *Difference-in-differences estimates of the effect of college creation on local growth at municipalities with high and low high-skill concentration for* $p = 17\%$



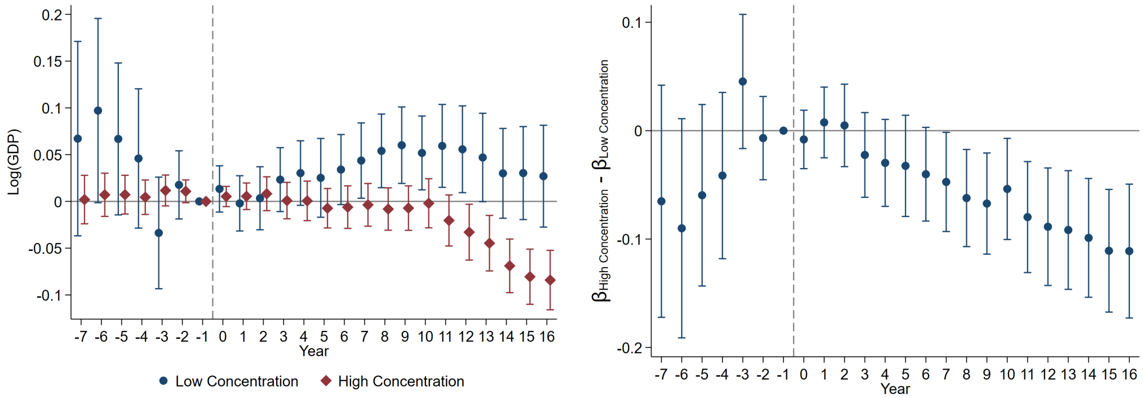

Note: High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Low (high) concentration municipalities are defined as those below (above) the $17^{th}$ percentile of high-skill concentration averaged within the three-year period pre-treatment, for those treated, or below (above) the same concentration level averaged in the initial two periods for the non-treated so as to get similar threshold levels in both cases. Vertical bars represent the 90% confidence interval.

Figure A16: *Difference-in-differences estimates of the effect of college creation on local growth at municipalities with high and low high-skill concentration controlling for multiple treatments*



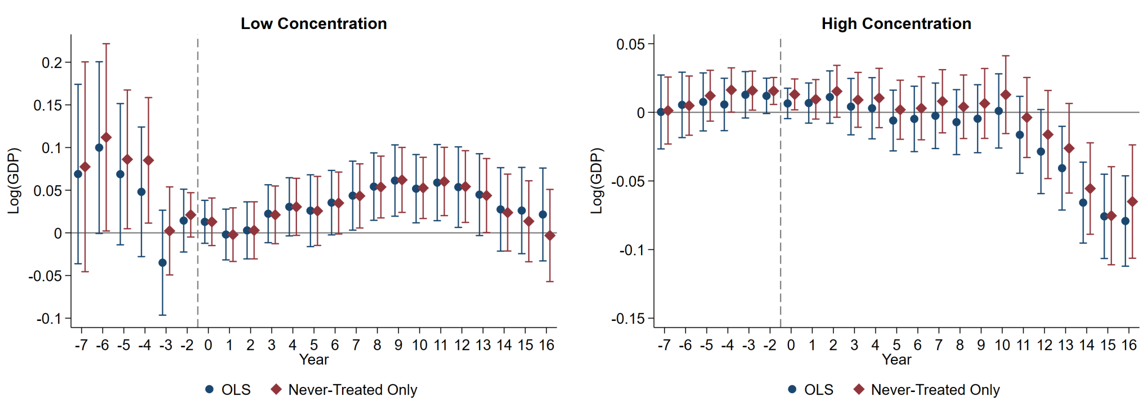

Note: Controls include the leads and lags of places treated twice and/or three times. High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Low (high) concentration municipalities are defined as those below (above) the $14^{th}$ percentile of high-skill concentration averaged within the three-year period pre-treatment, for those treated, or below (above) the same concentration level averaged in the initial two periods for the non-treated so as to get similar threshold levels in both cases. Vertical bars represent the 90% confidence interval.

Figure A17: *Difference-in-differences estimates of the effect of college creation on local growth at municipalities with high and low high-skill concentration (weighted by log(population))*



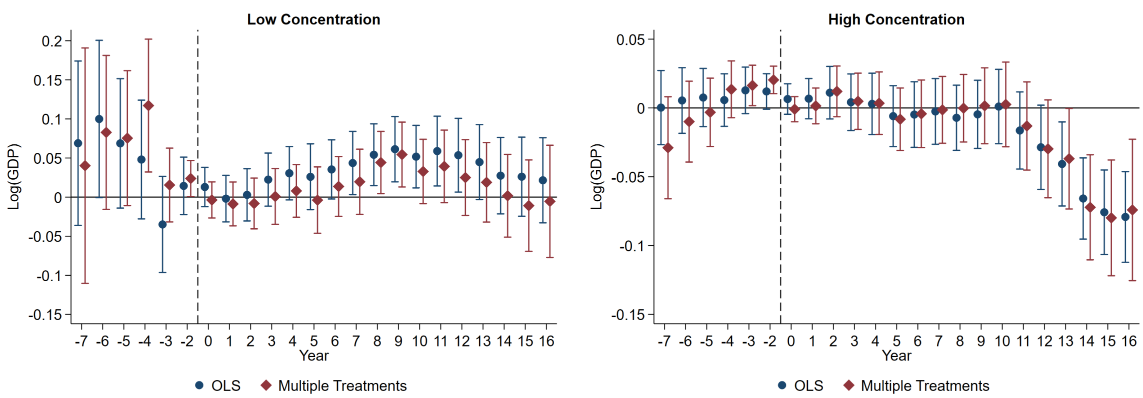

Note: High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Low (high) concentration municipalities are defined as those below (above) the $14^{th}$ percentile of high-skill concentration averaged within the three-year period pre-treatment, for those treated, or below (above) the same concentration level averaged in the initial two periods for the non-treated so as to get similar threshold levels in both cases. Specification runs a weighted regression using the logarithm of local population as weights. Vertical bars represent the 90% confidence interval.

Figure A18: *Difference-in-differences estimates of the effect of college creation on local growth using robust estimators*



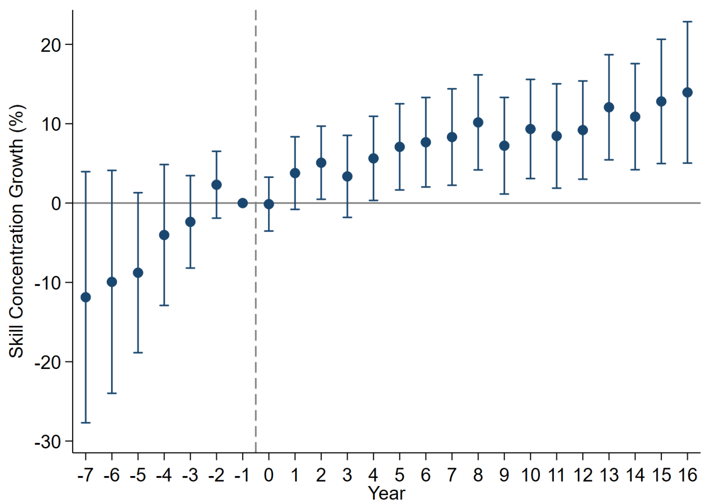

High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Low (high) concentration municipalities are defined as those below (above) the $14^{th}$ percentile of high-skill concentration averaged within the three-year period pre-treatment, for those treated, or below (above) the same concentration level averaged in the initial two periods for the non-treated so as to get similar threshold levels in both cases. Estimates restrict the control group to never-treated units and use the estimator proposed in Sun and Abraham (2021). Vertical bars represent the 90% confidence interval.

Figure A19: *Difference-in-differences estimates of the effect of college creation on local growth (baseline vs. non-binary treatment estimator)*



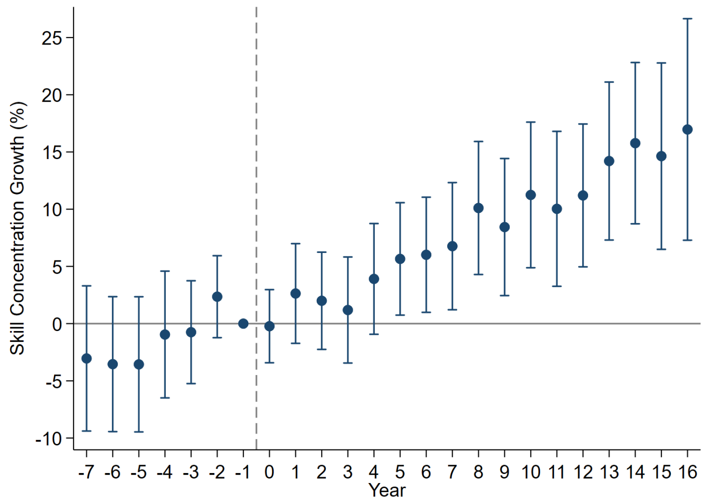

High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Low (high) concentration municipalities are defined as those below (above) the $14^{th}$ percentile of high-skill concentration averaged within the three-year period pre-treatment, for those treated, or below (above) the same concentration level averaged in the initial two periods for the non-treated so as to get similar threshold levels in both cases. Multiple Treatments estimates restrict the control group to never-treated units, only considers as treated the observations that saw an increase in the number of colleges, and use the estimator proposed in de Chaisemartin and D'Haultfœuille (2024). Vertical bars represent the 90% confidence interval.

Figure A20: *Difference-in-differences estimates of the effect of college creation on local high-skill concentration using the placebo group as control*



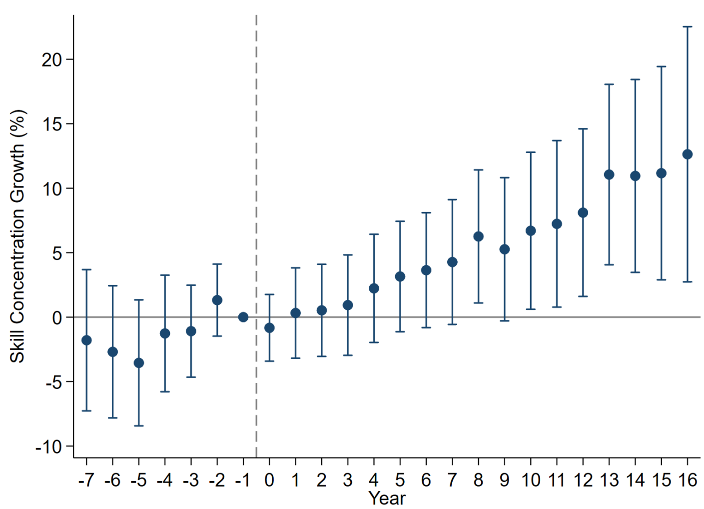

Note: High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Sample excludes observations with no workers, high-skill or not, at large firms. Sample match is on population level, share earning minimum wage or lower, share who only completed the $5^{th}$ grade, unemployment rate, and illiteracy rate, all in 2000. Treated observations are matched to those in control using the coarsened matching method in Iacus et al. (2012). Vertical bars represent the 95% confidence interval.

Figure A21: *Difference-in-differences estimates of the effect of college creation on local high-skill concentration for the no-college sample*



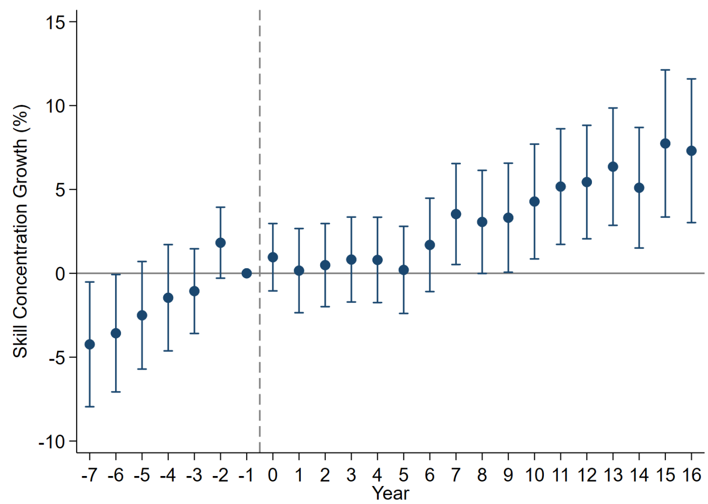

Note: High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Sample excludes observations with no workers, high-skill or not, at large firms. No-college sample only includes observations that do not have a college in all periods (control) or in the pre-treatment period (treated). Vertical bars represent the 95% confidence interval.

Figure A22: *Difference-in-differences estimates of the effect of college creation on local high-skill concentration controlling for multiple treatments*



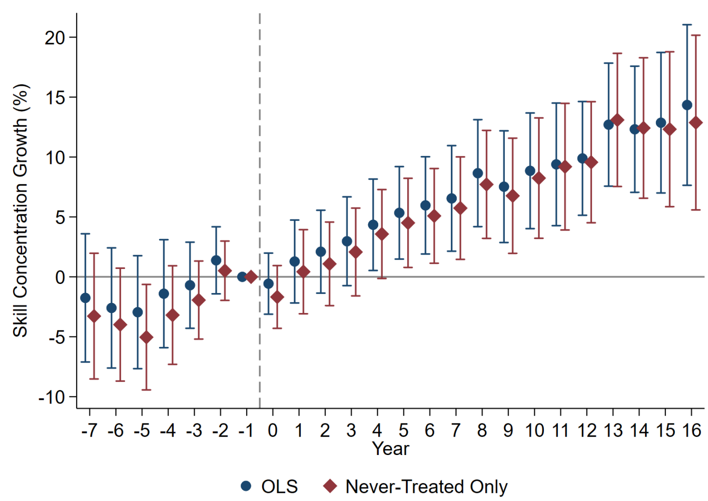

Note: High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Sample excludes observations with no workers, high-skill or not, at large firms. Controls include the leads and lags of places treated twice and/or three times. Vertical bars represent the 95% confidence interval.

Figure A23: *Difference-in-differences estimates of the effect of college creation on local high-skill concentration (HHI-based)*



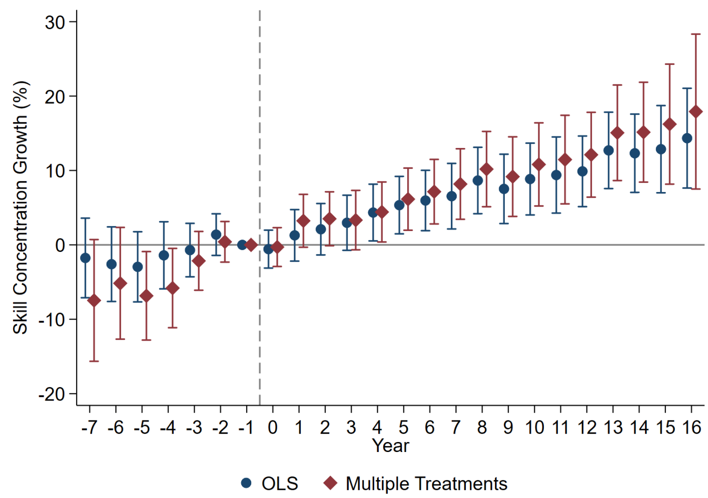

Note: High-skill concentration is the HHI-based measure calculated using firm size bins, i.e. for each firm-size range in the RAIS dataset I calculate the sum of the square of the corresponding local employment share. High-skill workers are those with at least some college education, though they might not have finished their degree. Sample excludes observations with no workers, high-skill or not, at large firms. Vertical bars represent the 95% confidence interval.

Figure A24: *Comparison between baseline and robust difference-in-difference estimates*



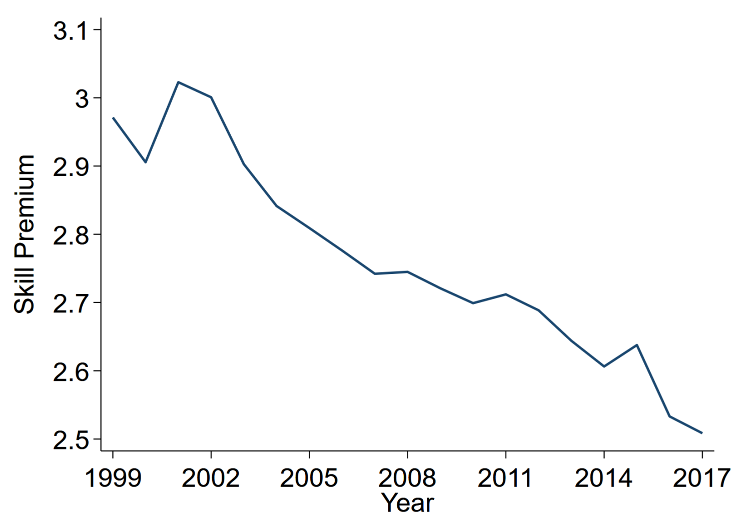

Note: High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Sample excludes observations with no workers, high-skill or not, at large firms. Never-Treated Only estimates restrict the control group to never-treated units and use the estimator proposed in Sun and Abraham (2021). Vertical bars represent the 95% confidence interval.

Figure A25: *Comparison between baseline results and estimates using non-binary treatment*



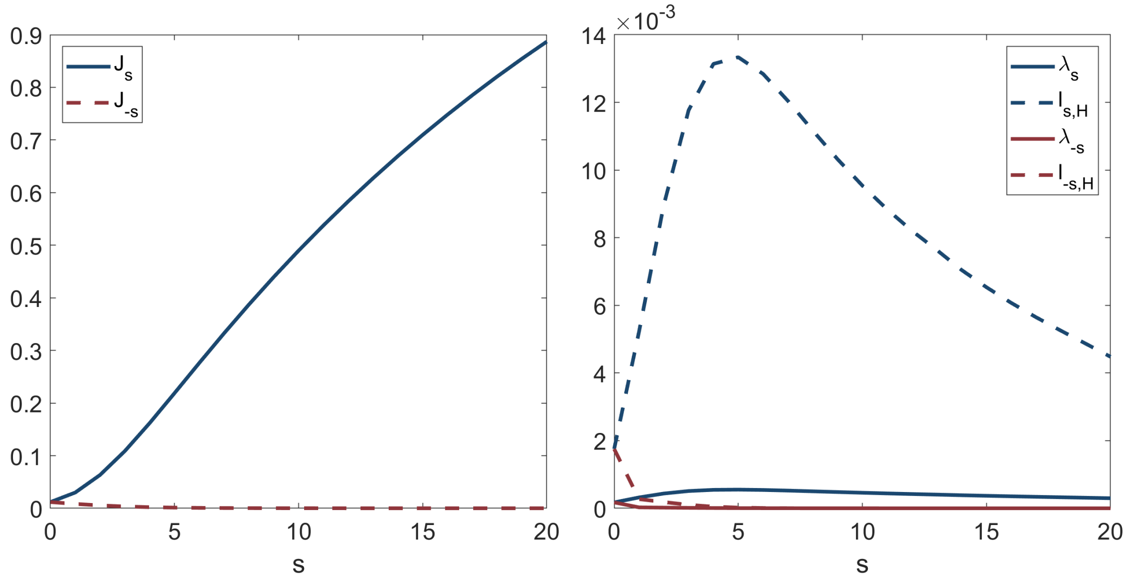

Note: High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Sample excludes observations with no workers, high-skill or not, at large firms. Multiple Treatments estimates restrict the control group to never-treated units, only considers as treated the observations that saw an increase in the number of colleges, and use the estimator proposed in de Chaisemartin and D'Haultfœuille (2024). Vertical bars represent the 95% confidence interval.

Figure A26: *Evolution of the skill premium in Brazil*




Note: Skill premium consists of the weighted average of the municipality-level ratio between high and low-skill wages.

Figure A27: *Left: Value function curves; Right: leader's and laggard's high-skill labor and investment choices, all as a function of the gap s*




Note: $J_s$ ($J_{-s}$) refers to the value function of the leader (follower). $\lambda_s$ ($\lambda_{-s}$) refers to R&D investment by the leader (follower). $l_{s,H}$ ($l_{-s,H}$) refers to high-skill labor hired by the leader (follower).

Figure A28: *Share of active R&D catch-up and skill concentration as a function of the gap s*

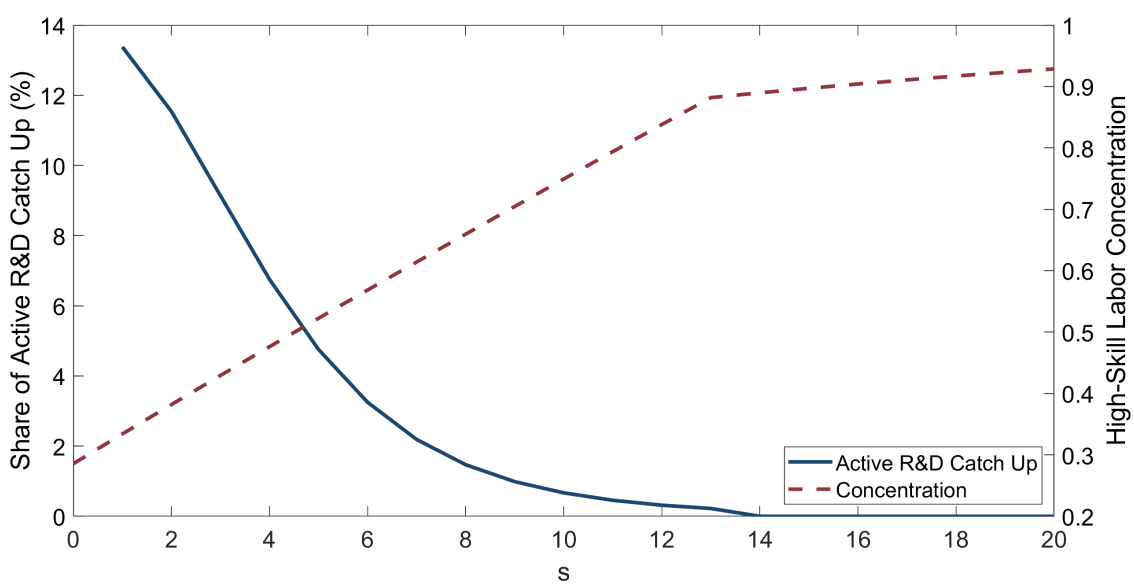


Figure A29: *Growth rate, wage premium, and high-skill concentration as a function of the gap (without the non-innovative, outside firm)*

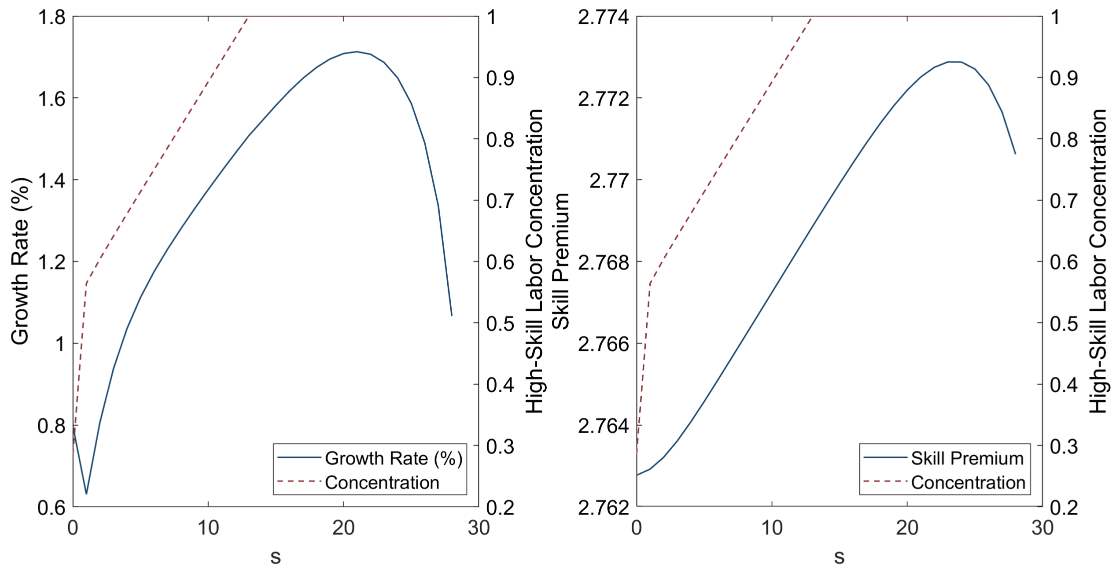



Figure A30: *Growth rate, wage premium, and high-skill concentration as a function of the gap for different values of the R&D labor elasticity*

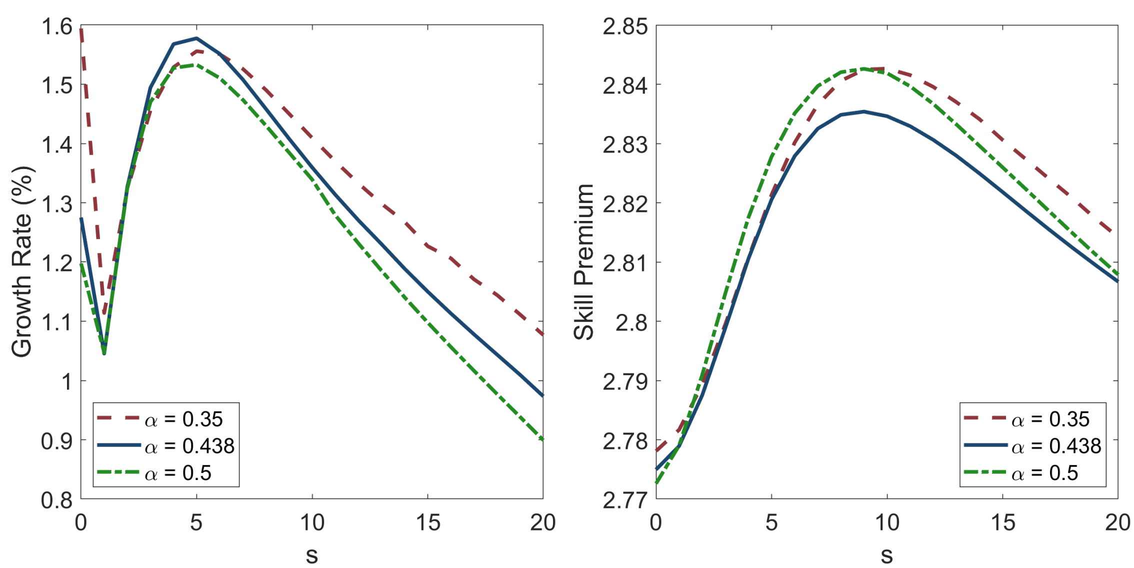
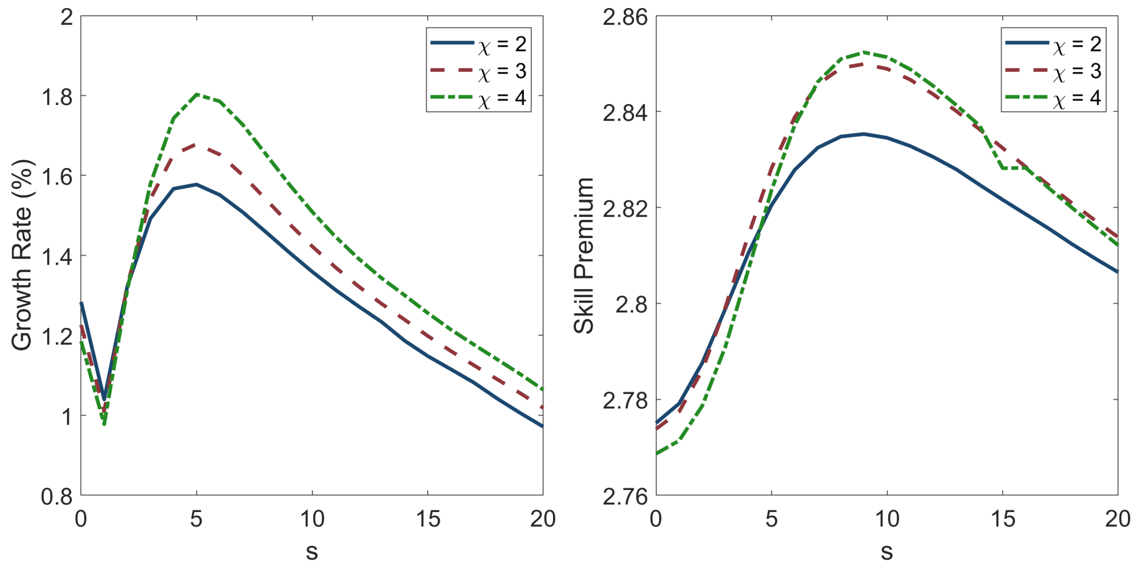
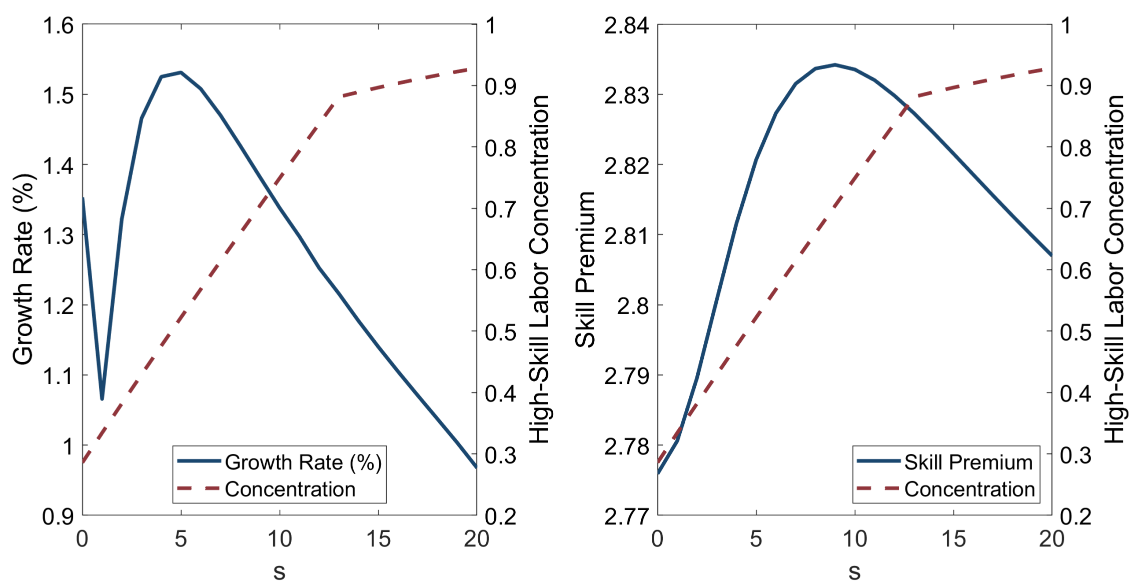


Figure A31: *Growth rate, wage premium, and high-skill concentration as a function of the gap for different convexity values of the R&D cost function*




Note: For each curve, we consider the following R&D cost function: $C(\lambda_s) = \rho \frac{\lambda_s^{\chi}}{\chi}$.

Figure A32: *Growth rate, wage premium, and high-skill concentration as a function of the gap using a smaller numerical adjustment*




Note: While in the baseline estimation I adjust aggregate demand by adding 0.005 to it, i.e. $D(t) = w_L(t)l_{i,L}(t) + w_o(t)l_o(t) + \pi(t) + 0.005$, in this plot I use 0.003 instead.

Figure A33: *Growth rate, wage premium, and high-skill concentration, calculated without the non-innovative sector, as a function of the gap*

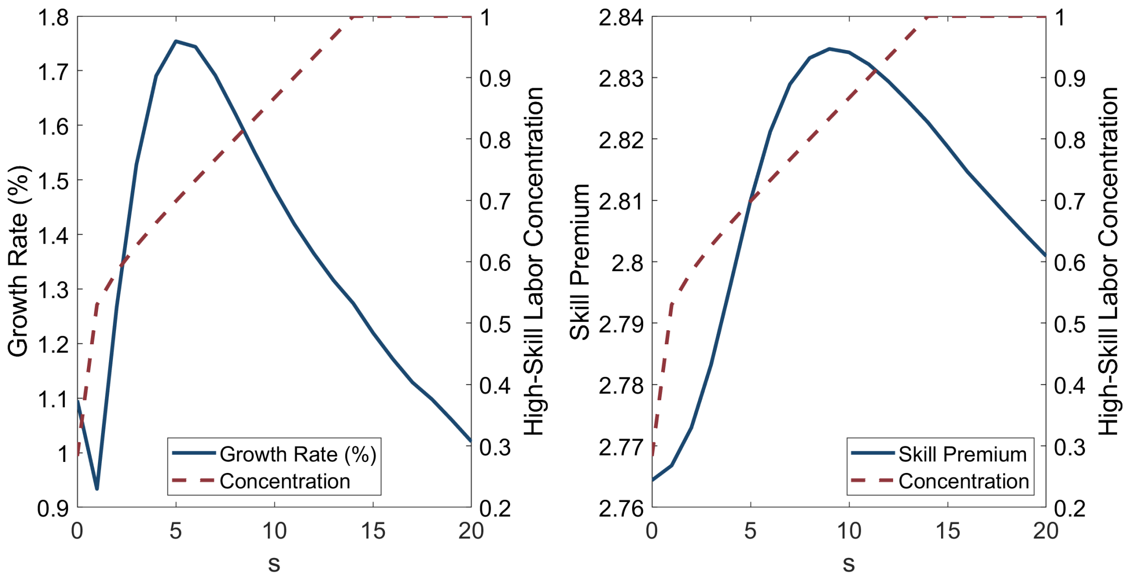



Figure A34: *Growth rate, wage premium, and high-skill concentration, calculated without the non-innovative sector, as a function of the gap*

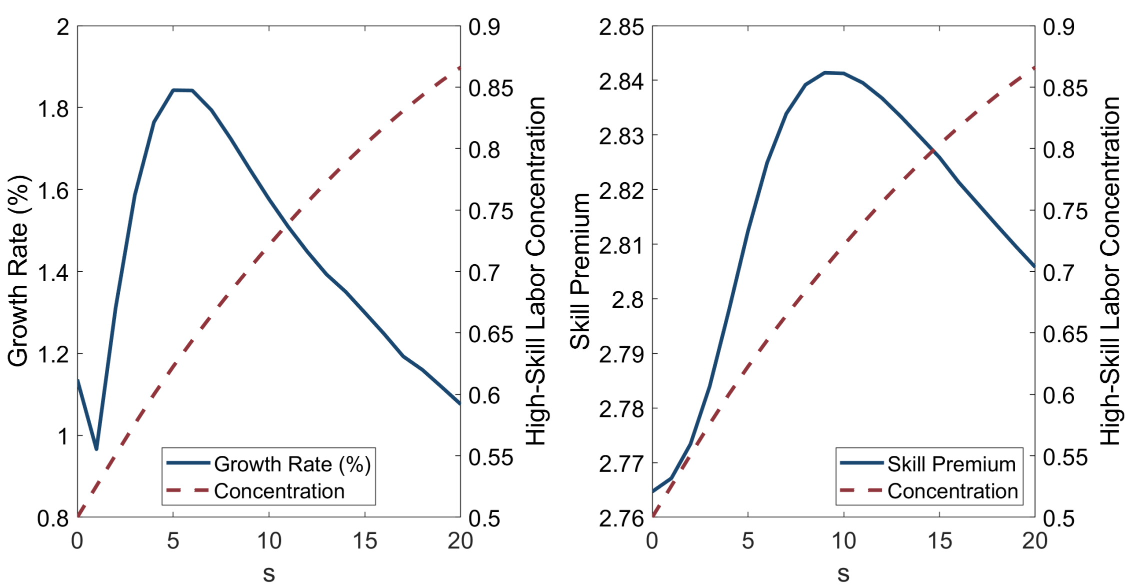
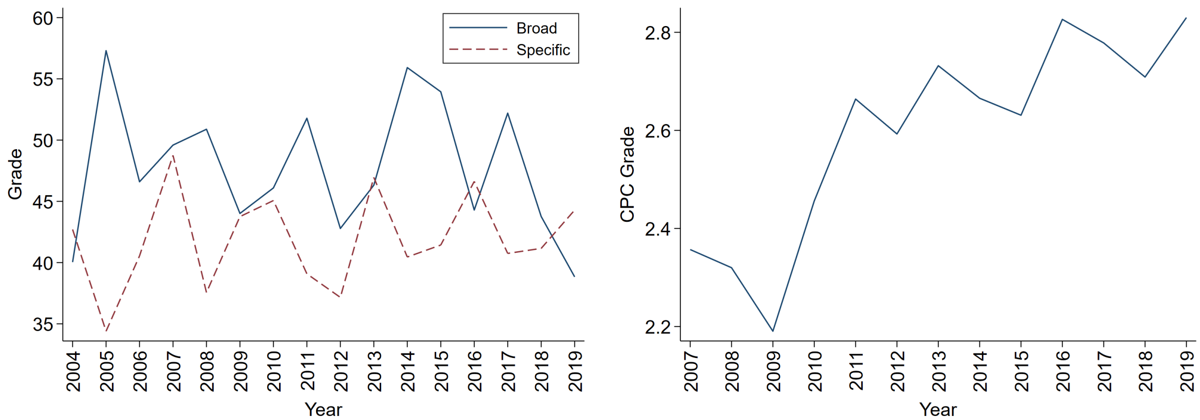
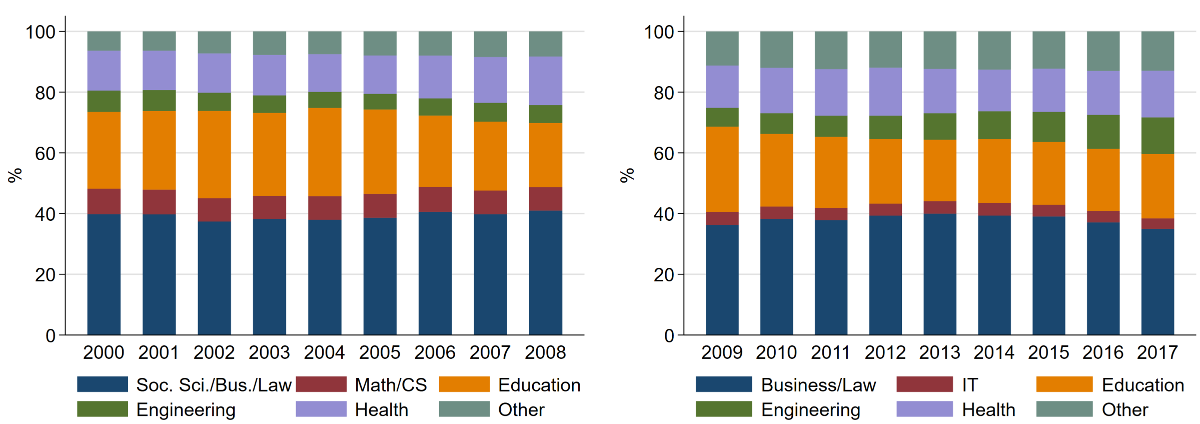


Figure A35: *Evolution of the government-run National Student Performance Exam in Brazil (ENADE) and the Preliminary Course Score (CPC)*




Note: Broad refers to the part of the exam that is common to all degrees. Specific refers to the part of the exam that is specific to a degree. CPC is a composite indicator of quality which takes into account the ENADE grade, teaching staff quality, student feedback, and an indicator of learning value added.

Figure A36: *Evolution of college graduates composition between areas of study*




Note: A new area of study classification 2009 onwards leads to a breakdown in the series.

Figure A37: *Growth rate as a function of the gap s and the distribution of gaps in the economy at different $L_H$*

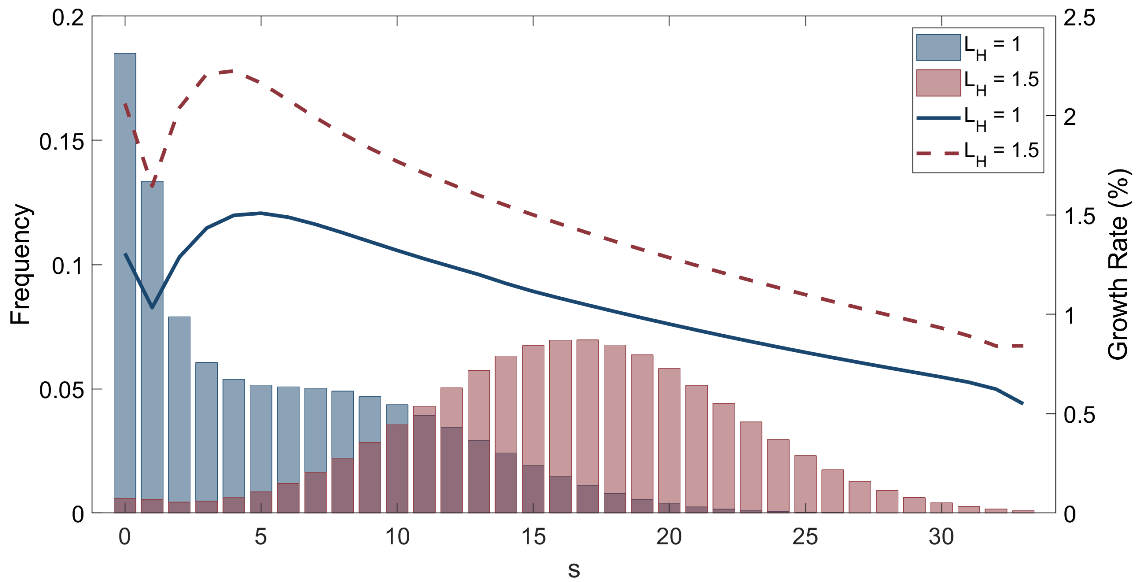
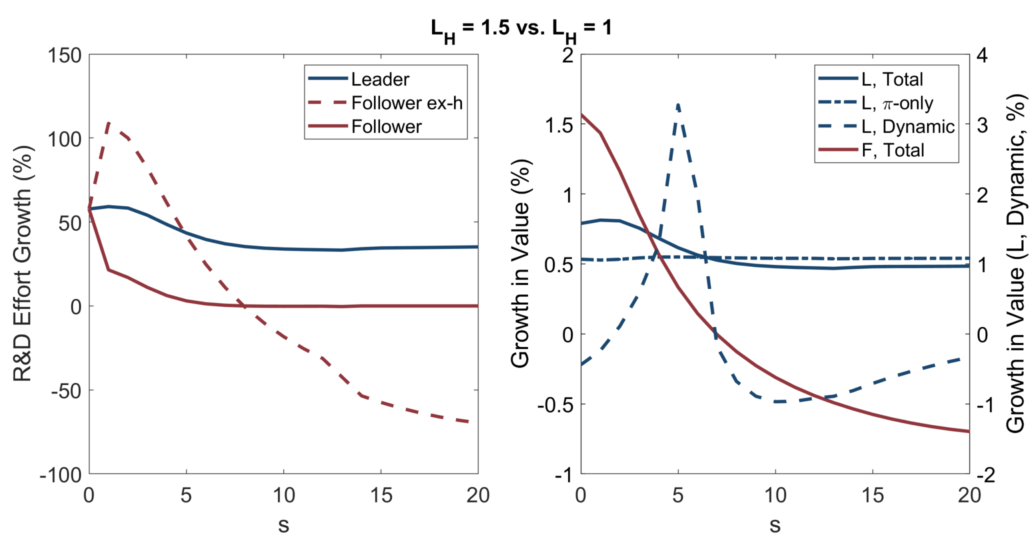
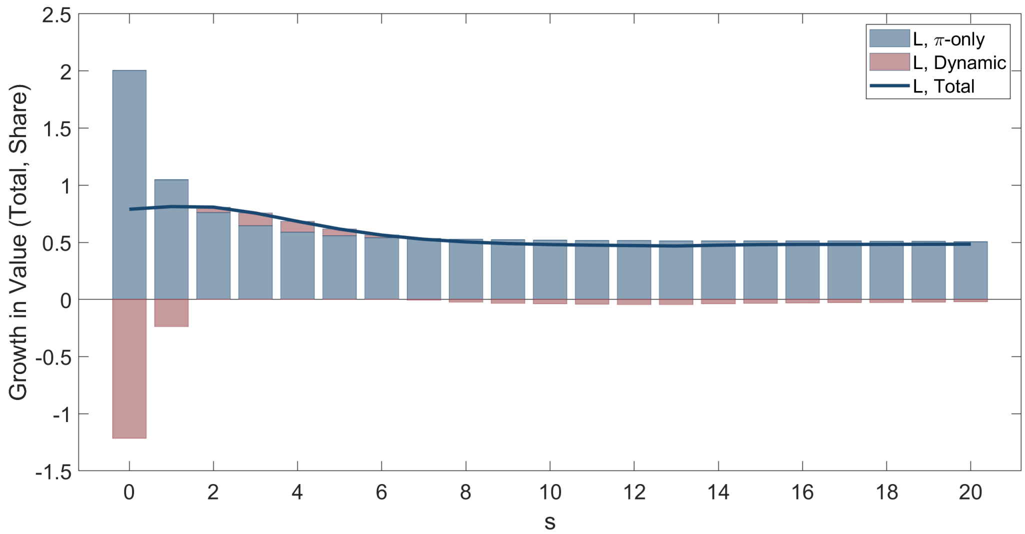



Figure A38: *R&D effort growth and breakdown of the change in the value function as high-skill supply increases from 1 to 1.5*



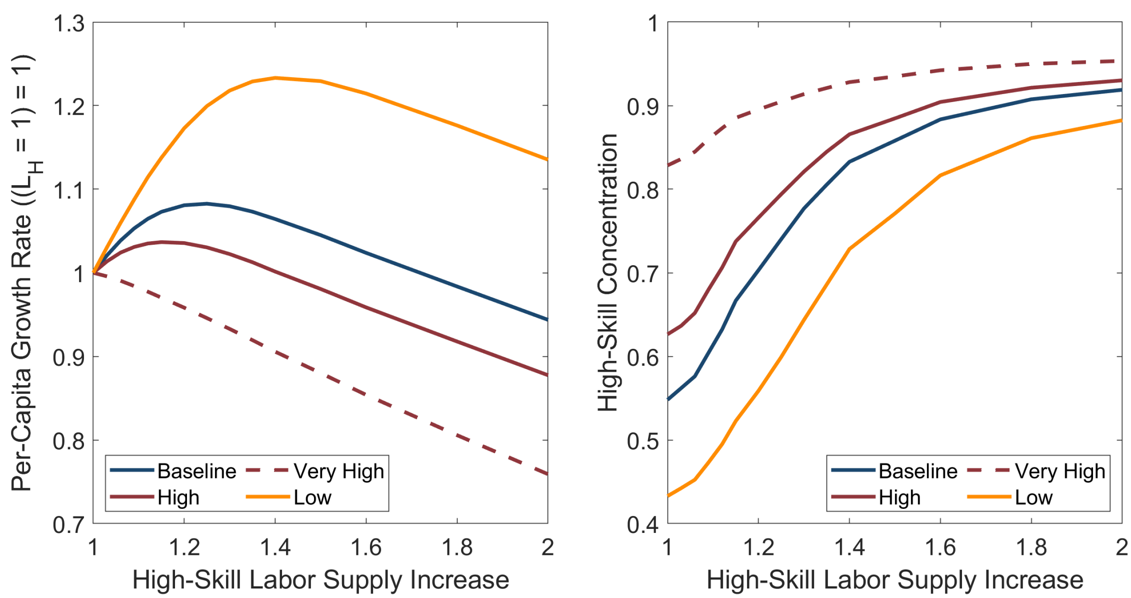

Note: Ex-h refers to the follower's R&D effort without the catch-up term. L, Total, $\pi$-only, and Dynamic refer to the change in the leader's total, profit-only, and dynamic parts of its value function, respectively. F, Total refers to the change in the follower's total value function.

Figure A39: *Breakdown of the change in the leader's value function as high-skill supply increases from 1 to 1.5*




Note: L, Total, $\pi$-only, and Dynamic refer to the change in the leader's total, profit-only, and dynamic parts of its value function, respectively.

Figure A40: *Growth and skill concentration as a function of human capital supply for different levels of initial skill concentration*




Note: Baseline refers to values using parameter estimates from Table 4. High, Very High, and Low (Skill Concentration) use the same set of parameters as the baseline scenario except for $h_c$ whose value is $h_{c,baseline}/1.05$, $h_{c,baseline}/1.25$, and $1.1h_{c,baseline}$, respectively.

Figure A41: *Effect of increasing high-skill labor supply on skill premium and high-skill unemployment for different increases in aggregate labor supply*

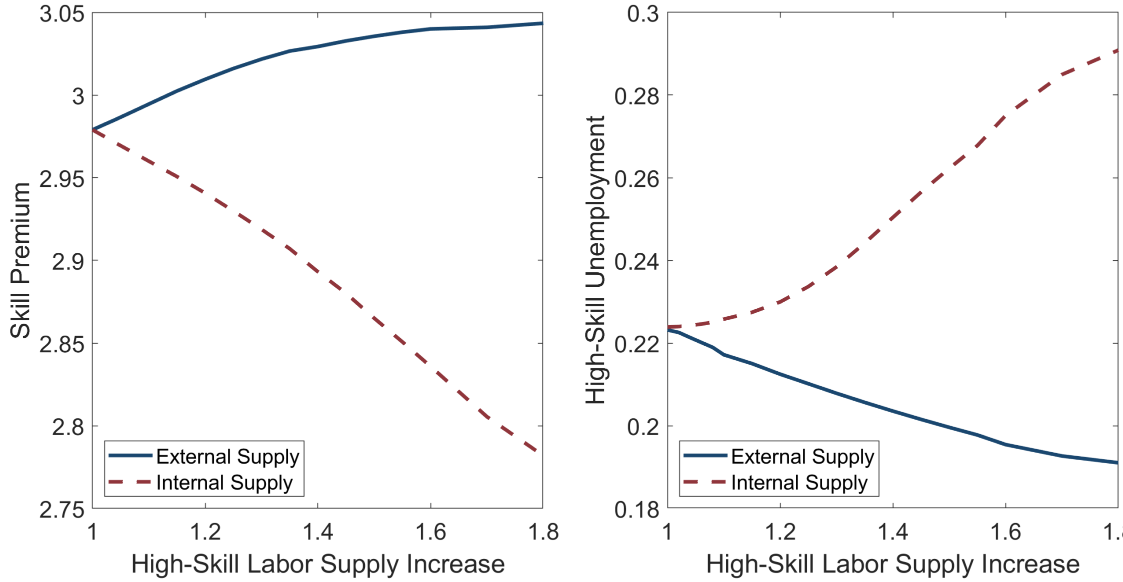
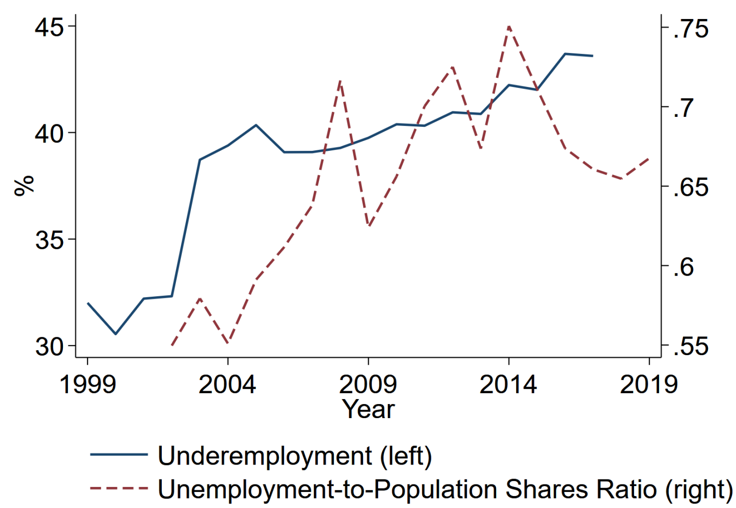



Figure A42: *Evolution of high-skill underemployment and unemployment-to-population shares ratio in Brazil*




Note: High-skill denotes those with 15+ years of study which corresponds to at least a college degree. Underemployment data comes from RAIS and is defined as an employee with a college degree working in Groups 3-9 in Brazil's occupational classification system (CBO). Unemployment and population shares come from the National Household Sample Survey (PNAD). Values for 2010 are interpolated as the household survey is not run during Census years.

Figure A43: *Active R&D catch-up as high-skill labor supply increases*

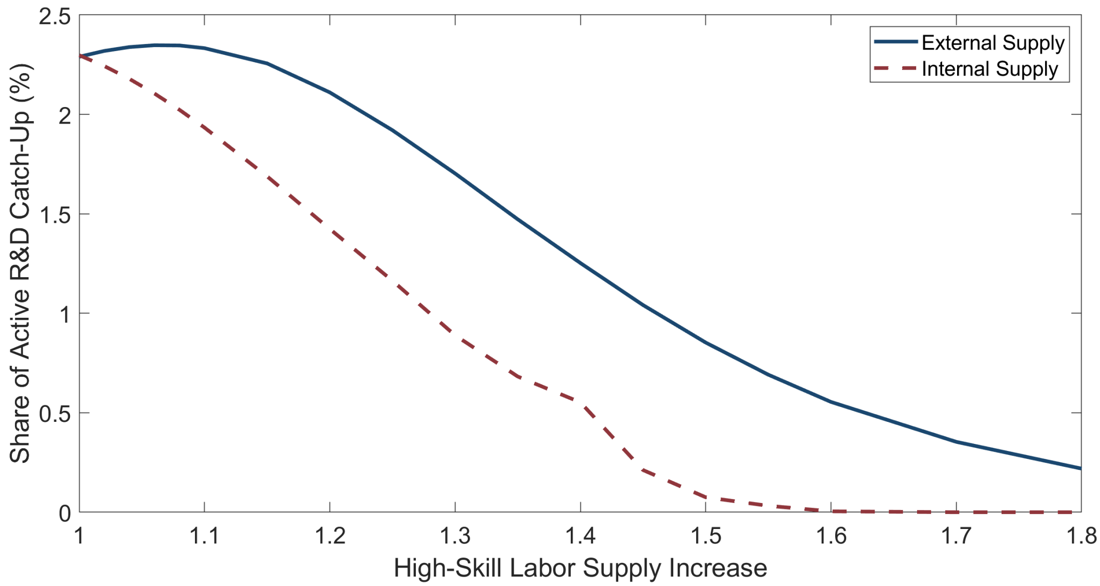



Figure A44: *Firms' R&D effort at* $L_H = 1.5$ *for the baseline and subsidy cases*

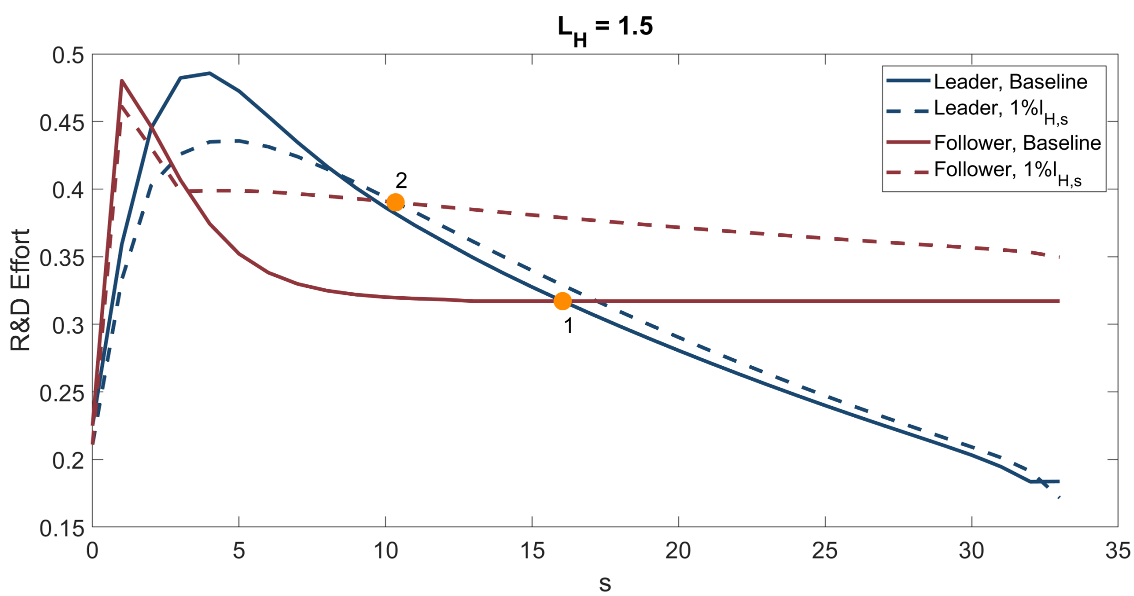


Table A.1: Summary statistics

| | Mean | St. Dev. | Obs | Min | Max |
|---|---|---|---|---|---|
| GDP Per Capita Growth | 0.031 | 0.181 | 74,209 | -0.840 | 12.7 |
| Skill Premium | 2.114 | 0.601 | 68,691 | 0.293 | 11.7 |
| Skill Premium - CT Workers | 2.086 | 0.682 | 68,171 | 0.355 | 11.8 |
| High-Skill Concentration | 0.624 | 0.266 | 74,209 | 0.000 | 1.0 |
| CT Worker Concentration | 0.608 | 0.287 | 73,733 | 0.000 | 1.0 |
| High-Skill Workers (th.) | 2.153 | 31.338 | 74,209 | 0.010 | 2319.5 |
| CT Workers (th.) | 1.188 | 15.337 | 74,209 | 0.000 | 1076.3 |
| Non-High-Skill Workers (th.) | 9.453 | 86.051 | 74,209 | 0.001 | 5768.9 |
| Electricity Consumption Growth | -0.013 | 0.243 | 74,209 | -0.942 | 7.4 |
| Real Wages (th.) | 1.543 | 0.478 | 74,209 | 0.232 | 9.4 |
| Population (mm.) | 0.036 | 0.209 | 74,209 | 0.001 | 12.0 |
| Total Workers (th.) | 11.299 | 114.655 | 74,209 | 0.003 | 8042.9 |
| Net New Workers Per Capita | 0.002 | 0.017 | 74,209 | -0.867 | 1.1 |
| Minimum Wage Population Share | 0.010 | 0.013 | 74,209 | 0.000 | 0.6 |
| High-Skill Population Share | 0.024 | 0.024 | 74,209 | 0.000 | 1.8 |
| CT Workers Population Share | 0.014 | 0.015 | 74,209 | 0.000 | 1.6 |
| Informality Share (2000) | 0.512 | 0.165 | 74,209 | 0.073 | 1.0 |

High-skill workers are those with at least some college education, though they might not have finished their degree. High-skill concentration is the local share of high-skill people working at large firms over total local supply. Skill premium is defined as the ratio between the average high-skill wage over the average non-high-skill wage within a municipality. CT (critical-thinking) workers are those with at least some college education who are also employed in occupations at the top skill quartile for one of the following: Math, Science, Critical Thinking, Active Learning, and Complex Problem Solving.

Table A.2: Summary statistics on municipalities

| | Full-Sample | | Matched | |
|---|---|---|---|---|
| | Control | Treatment | Control | Treatment |
| N | 4839.0 | 408.0 | 2081.0 | 385.0 |
| Population (th.) | 15.6 (91.3) | 97.3 (144.8) | 34.7 (76.6) | 81.4 (87.8) |
| Share Earning Min Wage (%) | 11.9 (19.2) | 3.4 (6.3) | 2.9 (4.9) | 2.8 (4.8) |
| Real Wage | 1213.2 (528.7) | 1610.1 (571.9) | 1558.0 (568.3) | 1613.7 (573.3) |
| Unemployment Rate (%) | 9.9 (5.9) | 13.3 (4.9) | 12.9 (4.9) | 13.0 (4.7) |
| Share Earning $<$ 0.25 x Min Wage (%) | 41.2 (22.3) | 23.3 (15.6) | 25.0 (16.6) | 22.7 (15.3) |
| Share in Agriculture (%) | 22.7 (23.7) | 11.5 (15.1) | 16.9 (19.4) | 11.4 (14.7) |
| Share in Industry (%) | 18.9 (19.4) | 25.5 (14.3) | 23.3 (18.0) | 25.8 (14.2) |
| Share in Services (%) | 66.3 (26.4) | 63.2 (17.4) | 61.7 (21.2) | 63.0 (17.2) |
| Illiterate Share (%) | 4.6 (7.0) | 2.5 (3.0) | 2.4 (3.3) | 2.4 (2.6) |
| $<$ 5$^{\text{th}}$ Grade Share (%) | 16.8 (12.6) | 10.4 (7.2) | 10.4 (7.2) | 10.4 (7.0) |
| = 5$^{\text{th}}$ Grade Share (%) | 16.8 (10.8) | 13.8 (6.9) | 17.1 (10.0) | 13.9 (6.5) |
| 6$^{\text{th}}$ to $<$ 9$^{\text{th}}$ Grade Share (%) | 15.2 (8.8) | 16.8 (5.5) | 17.3 (7.2) | 17.0 (5.5) |
| = 9$^{\text{th}}$ Grade Share (%) | 13.8 (9.5) | 17.2 (6.2) | 16.6 (7.9) | 17.4 (6.2) |
| Incomplete High-School Share (%) | 7.0 (5.6) | 9.6 (3.4) | 8.3 (4.6) | 9.7 (3.4) |
| High-School Share (%) | 22.4 (13.9) | 21.7 (8.4) | 20.6 (10.0) | 21.5 (8.4) |
| Incomplete College Share (%) | 2.5 (3.9) | 2.3 (1.7) | 2.4 (1.9) | 2.3 (1.7) |
| College+ Share (%) | 4.5 (4.0) | 5.7 (4.0) | 5.4 (3.4) | 5.6 (3.9) |

Table reports sample means and standard errors, the latter in parenthesis. Statistics are for the year 2000. Sample match is on population level, share earning minimum wage or lower, share who only completed the 5$^{th}$ grade, unemployment rate, and illiteracy rate, all in 2000. Treated observations are matched to those in control using the coarsened matching method in Iacus et al. (2012).

Table A.3: Summary statistics on pre-treated municipalities by skill concentration level

| | Treated Low Skill Concentration | Treated High Skill Concentration |
|---|---|---|
| N | 67.0 | 338.0 |
| Population (th.) | 81.1 (88.2) | 101.3 (153.9) |
| Share Earning Min Wage (%) | 3.9 (5.0) | 3.0 (6.2) |
| Real Wage | 1429.4 (393.0) | 1648.8 (595.2) |
| Unemployment Rate (%) | 12.5 (4.7) | 13.4 (5.0) |
| Share Earning $<$ 0.25 x Min Wage (%) | 25.0 (16.3) | 22.7 (15.3) |
| Share in Agriculture (%) | 10.1 (13.6) | 11.7 (15.2) |
| Share in Industry (%) | 26.5 (15.5) | 25.4 (14.1) |
| Share in Services (%) | 63.5 (18.0) | 63.1 (17.2) |
| Illiterate Share (%) | 2.7 (2.6) | 2.5 (3.1) |
| $<$ 5$^{\text{th}}$ Grade Share (%) | 10.7 (7.1) | 10.3 (7.1) |
| = 5$^{\text{th}}$ Grade Share (%) | 13.0 (7.4) | 13.9 (6.8) |
| 6$^{\text{th}}$ to $<$ 9$^{\text{th}}$ Grade Share (%) | 16.8 (6.8) | 16.8 (5.2) |
| = 9$^{\text{th}}$ Grade Share (%) | 16.8 (5.9) | 17.4 (6.2) |
| Incomplete High-School Share (%) | 9.8 (3.4) | 9.6 (3.4) |
| High-School Share (%) | 22.6 (10.6) | 21.5 (7.9) |
| Incomplete College Share (%) | 2.4 (1.5) | 2.3 (1.8) |
| College+ Share (%) | 5.1 (3.3) | 5.8 (4.1) |

Note: Table reports sample means and standard errors, the latter in parenthesis. Statistics are for the year 2000. High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. Low (high) concentration municipalities are defined as those below (above) the 14$^{th}$ percentile of high-skill concentration averaged within the three-year period pre-treatment, for those treated, or below (above) the same concentration level averaged in the initial two periods for the non-treated so as to get similar threshold levels in both cases.

Table A.4: Summary statistics for treated and last-treated municipalities

| | Control, Full-Sample | Treatment, Full-Sample |
|---|---|---|
| N | 32.0 | 133.0 |
| Population (th.) | 490.7 (1980.4) | 171.5 (414.8) |
| Share Earning Min Wage (%) | 7.1 (7.2) | 5.8 (6.9) |
| Real Wage | 1778.0 (530.9) | 1755.3 (468.2) |
| Unemployment Rate (%) | 7.4 (2.1) | 7.2 (2.8) |
| Share Earning $<$ 0.25 x Min Wage (%) | 18.9 (14.6) | 18.5 (14.7) |
| Share in Agriculture (%) | 6.9 (9.7) | 7.3 (12.0) |
| Share in Industry (%) | 26.4 (16.1) | 27.5 (17.7) |
| Share in Services (%) | 68.0 (18.3) | 65.3 (18.8) |
| Illiterate Share (%) | 0.6 (0.4) | 0.8 (0.8) |
| $< 5^{th}$ Grade Share (%) | 4.6 (3.3) | 5.7 (4.8) |
| $= 5^{th}$ Grade Share (%) | 5.9 (6.1) | 5.5 (3.0) |
| $6^{th}$ to $< 9^{th}$ Grade Share (%) | 8.4 (3.5) | 10.0 (5.1) |
| $= 9^{th}$ Grade Share (%) | 13.2 (4.6) | 13.9 (5.8) |
| Incomplete High-School Share (%) | 8.9 (3.3) | 9.1 (4.6) |
| High-School Share (%) | 42.7 (8.8) | 42.1 (10.4) |
| Incomplete College Share (%) | 3.0 (1.3) | 2.8 (1.4) |
| College+ Share (%) | 12.8 (7.0) | 10.1 (4.4) |

Table reports sample means and standard errors, the latter in parenthesis. Statistics are for the year 2010. Last-treated cohort receives treatment in 2019. Both groups consist of municipalities that have not yet been treated in 2010.

Table A.5: Summary statistics for not-yet-treated municipalities

| | Early Treated | Late Treated |
|---|---|---|
| N | 340.0 | 325.0 |
| Population (th.) | 103.9 (156.2) | 130.7 (602.4) |
| Share Earning Min Wage (%) | 3.0 (5.8) | 5.3 (9.9) |
| Real Wage | 1636.4 (587.3) | 1491.9 (532.8) |
| Unemployment Rate (%) | 13.3 (5.1) | 13.2 (4.9) |
| Share Earning $<$ 0.25 x Min Wage (%) | 22.2 (15.1) | 29.5 (18.9) |
| Share in Agriculture (%) | 12.2 (15.9) | 11.6 (16.3) |
| Share in Industry (%) | 26.2 (14.6) | 24.5 (16.0) |
| Share in Services (%) | 61.9 (17.7) | 64.9 (19.2) |
| Illiterate Share (%) | 2.5 (3.2) | 2.7 (2.7) |
| $< 5^{th}$ Grade Share (%) | 10.4 (7.4) | 12.1 (9.2) |
| $= 5^{th}$ Grade Share (%) | 14.0 (7.1) | 14.8 (7.6) |
| $6^{th}$ to $< 9^{th}$ Grade Share (%) | 16.9 (5.4) | 16.1 (6.9) |
| $= 9^{th}$ Grade Share (%) | 17.5 (6.3) | 15.6 (6.1) |
| Incomplete High-School Share (%) | 9.7 (3.5) | 8.6 (3.7) |
| High-School Share (%) | 21.2 (8.3) | 22.8 (9.9) |
| Incomplete College Share (%) | 2.3 (1.5) | 2.2 (1.9) |
| College+ Share (%) | 5.6 (3.4) | 5.2 (4.1) |

Table reports sample means and standard errors, the latter in parenthesis. Statistics are for the year 2000. Early Treated places consist of those that will be treated by 2005. Late Treated places consist of those that will be treated after 2005. In both cases, municipalities have not been treated yet.

Table A.6: Effect of large and small firm SSIV on high-skill and non-high-skill hiring, and energy consumption)

| | # High-Skill | | # CT Workers | | # Non-High-Skill | Energy Consumption |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| SSIV - Large Firms | 5.078*** | 0.542*** | 3.033*** | 0.594*** | -2.208*** | 0.0653 |
| | (0.263) | (0.068) | (0.146) | (0.047) | (0.316) | (0.146) |
| SSIV - Small Firms | 9.311*** | 7.413*** | 4.953*** | 4.296*** | 5.416* | -0.559 |
| | (0.845) | (0.469) | (0.400) | (0.219) | (2.214) | (0.531) |
| SSIV - Total | | | | | | -1.576*** |
| | | | | | | (0.236) |
| N | 74,090 | 74,090 | 73,684 | 73,684 | 74,090 | 74,090 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| $\mathbb{1}_{\{HS\ Conc_{i,t-1}>p\}}$ | 0 | 1 | 0 | 1 | | |

High-skill concentration is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. CT (critical-thinking) workers are those with at least some college education who are also employed in occupations at the top skill quartile for one of the following: Math, Science, Critical Thinking, Active Learning, and Complex Problem Solving. # refers to workers per capita. SSIV - Total refers to the SSIV constructed using loans to both small and large firms. Energy Consumption refers to the capital proxy variable calculated using the change in local electricity consumption and is winsorized at the $1^{st}$ and $99^{th}$ percentiles. All specifications control for the lagged local sum of shares $s_{i,t-2}$. Local-level controls (all lagged to be contemporaneous with the SSIV): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. $\mathbb{1}\{HSConc_{i,t-1} > p\}$ indicates whether the dependent variable and the SSIVs are interacted with a dummy for being below (0) or above (1) the high-skill concentration threshold which is set at the $16^{th}$ percentile. Municipality-clustered standard errors are in parentheses. $^{*}$, $^{**}$, $^{***}$ denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table A.7: Effect of high-skill concentration in large firms on local skill premium in places with high and low concentration

| | Skill Premium | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=0 \times HS\ Conc._{t-1}$ | 9.286** | 9.669** | 8.294** | 9.630** | 9.641** | 9.692** | 9.107** |
| | (2.964) | (2.989) | (2.936) | (3.174) | (2.980) | (3.024) | (3.107) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=1 \times HS\ Conc._{t-1}$ | -1.280** | -1.358** | -1.237** | -1.349* | -1.381** | -1.341** | -1.260* |
| | (0.471) | (0.478) | (0.477) | (0.533) | (0.481) | (0.482) | (0.534) |
| N | 68,582 | 68,582 | 68,582 | 68,582 | 68,582 | 68,582 | 68,582 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | | Yes | Yes | Yes | Yes | Yes | Yes |
| Informality | | | Yes | | | | |
| Non-High-Skill | | | | Yes | | | Yes |
| Capital Proxy | | | | | Yes | | Yes |
| Employment HHI | | | | | | Yes | |
| LIML | | | | | | | |
| Joint F-statistic | 42.9 | 42.1 | 37.1 | 14.9 | 17.6 | 31.9 | 6.7 |
| J-test, p-value | 0.51 | 0.96 | 0.08 | . | 0.55 | 0.60 | . |
| OP F-statistic, $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=0$ | 32.9 | 32.8 | 32.3 | 35.0 | 31.7 | 37.2 | 36.2 |
| OP Critical Value, $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=0$ | 17.5 | 17.5 | 17.2 | 23.1 | 16.4 | 23.1 | 23.1 |
| OP F-statistic, $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=1$ | 72.7 | 71.4 | 63.9 | 85.2 | 65.6 | 103.8 | 79.4 |
| OP Critical Value, $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=1$ | 10.6 | 10.6 | 9.2 | 23.1 | 8.5 | 23.1 | 23.1 |

High-skill concentration ($HS\ Conc$) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. High-skill concentration threshold $p$ is set at the $10^{th}$ percentile. Skill premium is defined as the ratio between the average high-skill wage over the average non-high-skill wage within a municipality and it is winsorized at the $1^{st}$ and $99^{th}$ percentiles. Columns (1)-(4) use $B_{i,t-2,large}$ and $B_{i,t-2,small}$ as instruments, the former interacted with $\mathbb{1}\{HSConc_{i,t-1} > p\}$ and the latter with $\mathbb{1}\{\{HSConc_{i,t-1} > p\} = 0\}$. Columns (5) and (6) add to the IV set a SSIV calculated using both small and large shocks together, and using total employment shares as exposure shares. All specifications control for the lagged local sum of shares $s_{i,t-2}$. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. Informality refers to the 2000 ratio of informal workers over total employment interacted with year fixed-effects. Non-High-Skill refers to the total non-high-skill hiring instrumented with the SSIVs. Capital Proxy refers to the capital proxy variable calculated using the change in local electricity consumption instrumented with the SSIVs. Employment HHI refers to the HHI measure of concentration calculated for total employment. J-test refers to the overidentification test in Hansen (1982). OP F-statistic and Critical Value refer, respectively, to the Olea-Pflueger effective F-statistic and the critical value for a 5% significance level and a 10% "worst-case" bias. Municipality-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table A.8: Summary statistics of shocks and shares

| | Mean | St. Dev. | IQR | Max | Obs |
|---|---|---|---|---|---|
| Shock - Large Firms | 0.908 | 2.216 | 1.102 | 7.794 | 594 |
| Shock - Large Firms - Residual | 0.000 | 1.117 | 0.530 | 8.214 | 594 |
| Share - Large Firms | 0.001 | 0.007 | 0.000 | 0.051 | 594 |
| Shock - Small Firms | 0.263 | 0.770 | 0.804 | 2.887 | 622 |
| Shock - Small Firms - Residual | 0.000 | 0.660 | 0.615 | 3.122 | 622 |
| Share - Small Firms | 0.001 | 0.002 | 0.001 | 0.011 | 622 |
| Effective Sample Size - Large | . | . | . | . | 26 |
| Effective Sample Size - Small | . | . | . | . | 190 |
| Number of Sectors | . | . | . | . | 60 |

Shocks consist of the yearly change at the national level of BNDES loans by sector and firm seize. Shares consist of the local-level lagged high-skill employment shares by sector and firm size. Shock statistics are weighted by the shares. Residual statistics refer to shocks residualized on year fixed-effects. The effective sample size is measured as the inverse renormalized Herfindahl index of the shares.

Table A.9: Effect of high-skill concentration in large firms on local GDP growth (polynomial regressors and instruments)

| | GDP Growth | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $HS\ Conc._{t-1}^3$ | -0.463*** | -0.513*** | -0.503*** | -0.524** | -0.455*** | -0.455*** | -0.410** | -0.400* |
| | (0.109) | (0.116) | (0.134) | (0.159) | (0.113) | (0.137) | (0.152) | (0.166) |
| $HS\ Conc._{t-1}$ | 0.523*** | 0.553*** | 0.546*** | 0.554*** | 0.539*** | 0.537*** | 0.531*** | 0.531*** |
| | (0.098) | (0.097) | (0.095) | (0.098) | (0.100) | (0.114) | (0.103) | (0.105) |
| Vertex | 0.61 | 0.60 | 0.60 | 0.59 | 0.63 | 0.63 | 0.66 | 0.67 |
| Std. Error | 0.06 | 0.06 | 0.06 | 0.08 | 0.06 | 0.06 | 0.10 | 0.11 |
| N | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Informality | | | Yes | | | | | |
| Non-High-Skill | | | | Yes | | | Yes | Yes |
| Capital Proxy | | | | | Yes | | Yes | Yes |
| Employment HHI | | | | | | Yes | | |
| LIML | | | | | | | | Yes |
| Joint F-statistic | 10.3 | 8.9 | 7.4 | 11.8 | 11.8 | 8.1 | 7.3 | 7.3 |
| J-test, p-value | 0.73 | 0.47 | 0.84 | 0.22 | 0.37 | 0.36 | 0.21 | 0.21 |

High-skill concentration (*HS Conc*) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. GDP Growth is real per-capita local GDP growth winsorized at the $1^{st}$ and $99^{th}$ percentiles. Columns (1)-(4) use $B_{i,t-2,large}$, $B^3_{i,t-2,large}$, $B_{i,t-2,small}$, and $B^3_{i,t-2,small}$ as instruments (winsorized at the $5^{th}$ and $95^{th}$ percentiles). Columns (5)-(7) add to the IV set a SSIV calculated using both small and large shocks together, and using total employment shares as exposure shares. Vertex refers to the point in the domain where the derivative with respect to the regressor of interest is zero (i.e. the point where the slope changes sign). Std. Error refers to the standard error of the vertex point-estimate. All specifications control for the lagged local sum of shares $s_{i,t-2}$. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. Informality refers to the 2000 ratio of informal workers over total employment interacted with year fixed-effects. Non-High-Skill refers to the total non-high-skill hiring instrumented with the SSIVs. Capital Proxy refers to the capital proxy variable calculated using the change in local electricity consumption instrumented with the SSIVs. Employment HHI refers to the HHI measure of concentration calculated for total employment. J-test refers to the overidentification test in Hansen (1982). Municipality-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table A.10: Effect of high-skill concentration in large firms on local GDP growth (polynomial instruments)

| | GDP Growth | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=0 × $HS\ Conc._{t-1}$ | 0.779*** | 0.803*** | 0.836*** | 0.836*** | 0.782*** | 0.758*** | 0.841*** | 0.861*** |
| | (0.181) | (0.184) | (0.186) | (0.192) | (0.184) | (0.188) | (0.190) | (0.198) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=1 × $HS\ Conc._{t-1}$ | -0.541*** | -0.611*** | -0.645*** | -0.540* | -0.547** | -0.511** | -0.401 | -0.375 |
| | (0.162) | (0.172) | (0.193) | (0.242) | (0.166) | (0.181) | (0.232) | (0.275) |
| N | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Informality | | | Yes | | | | | |
| Non-High-Skill | | | | Yes | | | Yes | Yes |
| Capital Proxy | | | | | Yes | | Yes | Yes |
| Employment HHI | | | | | | Yes | | |
| LIML | | | | | | | | Yes |
| Joint F-statistic | 21.0 | 19.3 | 15.1 | 11.7 | 15.7 | 14.5 | 7.6 | 7.6 |
| J-test, p-value | 0.36 | 0.14 | 0.49 | 0.05 | 0.14 | 0.13 | 0.10 | 0.10 |

High-skill concentration (*HS Conc*) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. High-skill concentration threshold $p$ is set at the 25$^{th}$ percentile. GDP Growth is real per-capita local GDP growth winsorized at the 1$^{st}$ and 99$^{th}$ percentiles. Columns (1)-(4) use $B_{i,t-2,large}$, $B^3_{i,t-2,large}$, $B_{i,t-2,small}$, and $B^3_{i,t-2,small}$ as instruments (winsorized at the 5$^{th}$ and 95$^{th}$ percentiles). Columns (5)-(7) add to the IV set a SSIV calculated using both small and large shocks together, and using total employment shares as exposure shares. Vertex refers to the point in the domain where the derivative with respect to the regressor of interest is zero (i.e. the point where the slope changes sign). Std. Error refers to the standard error of the vertex point-estimate. All specifications control for the lagged local sum of shares $s_{i,t-2}$. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. Informality refers to the 2000 ratio of informal workers over total employment interacted with year fixed-effects. Non-High-Skill refers to the total non-high-skill hiring instrumented with the SSIVs. Capital Proxy refers to the capital proxy variable calculated using the change in local electricity consumption instrumented with the SSIVs. Employment HHI refers to the HHI measure of concentration calculated for total employment. J-test refers to the overidentification test in Hansen (1982). Municipality-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table A.11: Effect of concentration of critical thinking workers in large firms on local GDP growth in places with high and low concentration

| | GDP Growth | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=0 × $CT\ Conc._{t-1}$ | 1.683*** | 1.580*** | 1.709*** | 2.030*** | 1.577*** | 1.589*** | 2.464** | 2.486** |
| | (0.475) | (0.475) | (0.503) | (0.612) | (0.476) | (0.474) | (0.768) | (0.779) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=1 × $CT\ Conc._{t-1}$ | -0.354** | -0.339** | -0.356** | -0.395** | -0.340** | -0.343** | -0.485** | -0.489** |
| | (0.125) | (0.125) | (0.132) | (0.141) | (0.126) | (0.126) | (0.176) | (0.178) |
| N | 73,684 | 73,684 | 73,684 | 73,684 | 73,684 | 73,684 | 73,684 | 73,684 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Informality | | | Yes | | | | | |
| Non-High-Skill | | | | Yes | | | Yes | Yes |
| Capital Proxy | | | | | Yes | | Yes | Yes |
| Employment HHI | | | | | | Yes | | |
| LIML | | | | | | | | Yes |
| Joint F-statistic | 24.1 | 22.0 | 25.8 | 13.6 | 15.1 | 17.3 | 4.1 | 4.1 |
| J-test, p-value | 0.20 | 0.08 | 0.21 | 0.08 | 0.08 | 0.09 | 0.46 | 0.46 |

High-skill concentration ($HS\ Conc$) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. High-skill concentration threshold $p$ is set at the $16^{th}$ percentile. GDP Growth is real per-capita local GDP growth winsorized at the $1^{st}$ and $99^{th}$ percentiles. Columns (1)-(4) use $B_{i,t-2,large}$ and $B_{i,t-2,small}$ as instruments, each interacted with $\mathbb{1}\{HSConc_{i,t-1} > p\}$. Columns (5)-(7) add to the IV set a SSIV calculated using both small and large shocks together, and using total employment shares as exposure shares. All specifications control for the lagged local sum of shares $s_{i,t-2}$. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. Informality refers to the 2000 ratio of informal workers over total employment interacted with year fixed-effects. Non-High-Skill refers to the total non-high-skill hiring instrumented with the SSIVs. Capital Proxy refers to the capital proxy variable calculated using the change in local electricity consumption instrumented with the SSIVs. Employment HHI refers to the HHI measure of concentration calculated for total employment. J-test refers to the overidentification test in Hansen (1982). Municipality-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table A.12: Effect of concentration of critical thinking workers in large firms on local skill premium in places with high and low concentration

| | Skill Premium | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=0 × $HS\ Conc._{t-1}$ | 11.80** | 12.04** | 9.527* | 13.38** | 12.05** | 12.20** | 12.71** |
| | (3.766) | (3.792) | (3.732) | (4.132) | (3.780) | (3.863) | (4.078) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=1 × $HS\ Conc._{t-1}$ | -1.918*** | -1.983*** | -1.720** | -2.289*** | -2.027*** | -1.951** | -2.177** |
| | (0.575) | (0.583) | (0.578) | (0.681) | (0.588) | (0.596) | (0.680) |
| N | 68,453 | 68,453 | 68,453 | 68,453 | 68,453 | 68,453 | 68,453 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | | Yes | Yes | Yes | Yes | Yes | Yes |
| Informality | | | Yes | | | | |
| Non-High-Skill | | | | Yes | | | Yes |
| Capital Proxy | | | | | Yes | | Yes |
| Employment HHI | | | | | | Yes | |
| LIML | | | | | | | |
| Joint F-statistic | 41.6 | 40.9 | 35.4 | 14.9 | 17.5 | 27.1 | 6.6 |
| J-test, p-value | 0.13 | 0.18 | 0.53 | . | 0.58 | 0.46 | . |

High-skill concentration ($HS\ Conc$) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. High-skill concentration threshold $p$ is set at the $10^{th}$ percentile. Skill premium is defined as the ratio between the average high-skill wage over the average wage of workers without any college education and is winsorized at the $1^{st}$ and $99^{th}$ percentiles. Columns (1)-(4) use $B_{i,t-2,large}$ and $B_{i,t-2,small}$ as instruments, the former interacted with $\mathbb{1}\{HSConc_{i,t-1} > p\}$ and the latter with $\mathbb{1}\{\{HSConc_{i,t-1} > p\} = 0\}$. Columns (5) and (6) add to the IV set a SSIV calculated using both small and large shocks together, and using total employment shares as exposure shares. All specifications control for the lagged local sum of shares $s_{i,t-2}$ and are weighted by the twice lagged log of local population. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. Informality refers to the 2000 ratio of informal workers over total employment interacted with year fixed-effects. Non-High-Skill refers to the total non-high-skill hiring instrumented with the SSIVs. Capital Proxy refers to the capital proxy variable calculated using the change in local electricity consumption instrumented with the SSIVs. Employment HHI refers to the HHI measure of concentration calculated for total employment. J-test refers to the overidentification test in Hansen (1982). Municipality-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table A.13: Effect of high-skill concentration in large firms on local skill premium in places with low and mid-level concentration

| | Skill Premium | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}<p_1\}}$=1 $\times$ $HS\ Conc._{t-1}$ | 9.878*** | 10.47*** | 10.75*** | 11.01*** | 9.412** |
| | (2.877) | (2.915) | (2.912) | (2.942) | (2.926) |
| $\mathbb{1}_{\{p_1<HS\ Conc._{t-1}<p_2\}}$=1 $\times$ $HS\ Conc._{t-1}$ | -4.960** | -3.774*** | -2.868*** | -2.493*** | -2.762*** |
| | (1.733) | (1.098) | (0.795) | (0.702) | (0.681) |
| Top Threshold $p_2$ | 30% | 40% | 50% | 60% | 60% |
| N | 68,582 | 68,582 | 68,582 | 68,582 | 68,582 |
| Time FE | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes |
| # IVs | 4 | 4 | 4 | 4 | 3 |
| Joint F-statistic | 19.2 | 24.7 | 36.1 | 38.3 | 55.1 |
| J-test, p-value | 0.10 | 0.08 | 0.04 | 0.01 | 0.21 |

High-skill concentration ($HS\ Conc$) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. High-skill concentration threshold $p_1$ is set at the $10^{th}$ percentile. Skill premium is defined as the ratio between the average high-skill wage over the average non-high-skill wage within a municipality and is winsorized at the $1^{st}$ and $99^{th}$ percentiles. Columns (1)-(4) use $B_{i,t-2,large}$ and $B_{i,t-2,small}$ as instruments, each interacted with $(\mathbb{1}\{HSConc_{i,t-1} < p_1\} = 1)$ and $(\mathbb{1}\{p_1 < HSConc_{i,t-1} < p_2\} = 1)$. Column (5) removes $B_{i,t-2,small}(\mathbb{1}\{p_1 < HSConc_{i,t-1} < p_2\} = 1)$ from the set of instruments. All specifications control for the lagged local sum of shares $s_{i,t-2}$. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. J-test refers to the overidentification test in Hansen (1982). Municipality-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table A.14: Effect of high-skill concentration in large firms on local skill premium in places with high and low concentration (non-interacted $B_{i,t-2,small}$)

| | Skill Premium | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=0 $\times$ $HS\ Conc._{t-1}$ | 13.73*** | 13.85*** | 13.86*** | 14.00*** | 13.83*** | 13.82*** | 13.95*** |
| | (2.532) | (2.503) | (2.641) | (3.033) | (2.494) | (2.500) | (3.390) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=1 $\times$ $HS\ Conc._{t-1}$ | -2.347** | -2.422** | -2.434** | -2.456* | -2.418** | -2.412** | -2.446* |
| | (0.903) | (0.904) | (0.931) | (0.997) | (0.903) | (0.908) | (1.064) |
| N | 68,582 | 68,582 | 68,582 | 68,582 | 68,582 | 68,582 | 68,582 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | | Yes | Yes | Yes | Yes | Yes | Yes |
| Informality | | | Yes | | | | |
| Non-High-Skill | | | | Yes | | | Yes |
| Capital Proxy | | | | | Yes | | Yes |
| Employment HHI | | | | | | Yes | |
| LIML | | | | | | | |
| Joint F-statistic | 16.5 | 17.1 | 15.2 | 11.3 | 17.5 | 15.0 | 4.7 |
| J-test, p-value | 0.66 | 0.92 | 0.18 | . | 0.95 | 0.95 | . |

High-skill concentration ($HS\ Conc$) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. High-skill concentration threshold $p$ is set at the $15^{th}$ percentile. Skill premium is defined as the ratio between the average high-skill wage over the average non-high-skill wage within a municipality and it is winsorized at the $1^{st}$ and $99^{th}$ percentiles. Columns (1)-(4) use $B_{i,t-2,large}$ and $B_{i,t-2,small}$ as instruments, the former interacted with $\mathbb{1}\{HSConc_{i,t-1} > p\}$ and the latter with $\mathbb{1}\{\{HSConc_{i,t-1} > p\} = 0\}$. Columns (5) and (6) add to the IV set a SSIV calculated using both small and large shocks together, and using total employment shares as exposure shares. All specifications control for the lagged local sum of shares $s_{i,t-2}$. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. Informality refers to the 2000 ratio of informal workers over total employment interacted with year fixed-effects. Non-High-Skill refers to the total non-high-skill hiring instrumented with the SSIVs. Capital Proxy refers to the capital proxy variable calculated using the change in local electricity consumption instrumented with the SSIVs. Employment HHI refers to the HHI measure of concentration calculated for total employment. J-test refers to the overidentification test in Hansen (1982). Municipality-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table A.15: Effect of high-skill concentration in large firms on local GDP growth in places with high and low concentration (non-tradables only)

| | GDP Growth | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=0 × $HS\ Conc._{t-1}$ | 1.270*** | 1.207*** | 1.259*** | 1.244*** | 1.209*** | 1.213*** | 1.290*** | 1.385*** |
| | (0.276) | (0.274) | (0.280) | (0.287) | (0.271) | (0.271) | (0.316) | (0.382) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=1 × $HS\ Conc._{t-1}$ | -0.576** | -0.561** | -0.583** | -0.571** | -0.559** | -0.565** | -0.578** | -0.615** |
| | (0.187) | (0.184) | (0.184) | (0.186) | (0.182) | (0.180) | (0.188) | (0.205) |
| N | 73,879 | 73,879 | 73,879 | 73,879 | 73,879 | 73,879 | 73,879 | 73,879 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Informality | | | Yes | | | | | |
| Non-High-Skill | | | | Yes | | | Yes | Yes |
| Capital Proxy | | | | | Yes | | Yes | Yes |
| Employment HHI | | | | | | Yes | | |
| LIML | | | | | | | | Yes |
| Joint F-statistic | 16.3 | 16.0 | 16.5 | 11.7 | 11.7 | 14.4 | 4.8 | 4.8 |
| J-test, p-value | 0.23 | 0.15 | 0.20 | 0.06 | 0.15 | 0.16 | 0.07 | 0.08 |

High-skill concentration ($HS\ Conc$) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. High-skill concentration threshold $p$ is set at the $16^{th}$ percentile. GDP Growth is real per-capita local non-tradables GDP growth winsorized at the $1^{st}$ and $99^{th}$ percentiles. Columns (1)-(4) use $B_{i,t-2,large}$ and $B_{i,t-2,small}$ as instruments, each interacted with $\mathbb{1}\{HSConc_{i,t-1} > p\}$. Columns (5)-(7) add to the IV set a SSIV calculated using both small and large shocks together, and using total employment shares as exposure shares. All specifications control for the lagged local sum of shares $s_{i,t-2}$. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. Informality refers to the 2000 ratio of informal workers over total employment interacted with year fixed-effects. Non-High-Skill refers to the total non-high-skill hiring instrumented with the SSIVs. Capital Proxy refers to the capital proxy variable calculated using the change in local electricity consumption instrumented with the SSIVs. Employment HHI refers to the HHI measure of concentration calculated for total employment. J-test refers to the overidentification test in Hansen (1982). Municipality-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table A.16: Effect of high-skill concentration in large firms on local skill premium in places with high and low concentration (non-tradables only)

| | Skill Premium | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=0 $\times$ $HS\ Conc._{t-1}$ | 11.51*** | 11.57*** | 10.77** | 11.60*** | 10.60** | 11.69*** | 10.37** |
| | (3.376) | (3.387) | (3.273) | (3.402) | (3.253) | (3.503) | (3.279) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=1 $\times$ $HS\ Conc._{t-1}$ | -2.366* | -2.422* | -2.322* | -2.447* | -2.381* | -2.129 | -2.238 |
| | (1.151) | (1.153) | (1.098) | (1.211) | (1.092) | (1.213) | (1.148) |
| N | 68,466 | 68,466 | 68,466 | 68,466 | 68,466 | 68,466 | 68,466 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | | Yes | Yes | Yes | Yes | Yes | Yes |
| Informality | | | Yes | | | | |
| Non-High-Skill | | | | Yes | | | Yes |
| Capital Proxy | | | | | Yes | | Yes |
| Employment HHI | | | | | | Yes | |
| LIML | | | | | | | |
| Joint F-statistic | 16.7 | 16.1 | 15.2 | 14.2 | 13.8 | 13.6 | 7.4 |
| J-test, p-value | 0.67 | 0.92 | 0.17 | . | 0.63 | 0.17 | . |

High-skill concentration ($HS\ Conc$) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. High-skill concentration threshold $p$ is set at the $13^{th}$ percentile. Skill premium is defined as the ratio between the average high-skill wage over the average non-high-skill wage within a municipality and it is winsorized at the $1^{st}$ and $99^{th}$ percentiles. Columns (1)-(4) use $B_{i,t-2,large}$ and $B_{i,t-2,small}$ as instruments, the former interacted with $\mathbb{1}\{HSConc_{i,t-1} > p\}$ and the latter with $\mathbb{1}\{\{HSConc_{i,t-1} > p\} = 0\}$. Columns (5) and (6) add to the IV set a SSIV calculated using both small and large shocks together, and using total employment shares as exposure shares. All specifications control for the lagged local sum of shares $s_{i,t-2}$ and are weighted by the twice lagged log of local population. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. Informality refers to the 2000 ratio of informal workers over total employment interacted with year fixed-effects. Non-High-Skill refers to the total non-high-skill hiring instrumented with the SSIVs. Capital Proxy refers to the capital proxy variable calculated using the change in local electricity consumption instrumented with the SSIVs. Employment HHI refers to the HHI measure of concentration calculated for total employment. J-test refers to the overidentification test in Hansen (1982). Municipality-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table A.17: Effect of high-skill concentration in large firms on local GDP growth in places with high and low concentration (weighted by log of population)

| | GDP Growth | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=0 × $HS\ Conc._{t-1}$ | 0.813** | 0.705** | 0.855** | 1.044** | 0.688* | 0.732* | 1.199** | 1.196** |
| | (0.259) | (0.261) | (0.284) | (0.330) | (0.268) | (0.290) | (0.401) | (0.400) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=1 × $HS\ Conc._{t-1}$ | -0.559** | -0.565** | -0.600*** | -0.681*** | -0.554** | -0.564** | -0.716** | -0.715** |
| | (0.180) | (0.177) | (0.182) | (0.201) | (0.177) | (0.182) | (0.219) | (0.218) |
| N | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Informality | | | Yes | | | | | |
| Non-High-Skill | | | | Yes | | | Yes | Yes |
| Capital Proxy | | | | | Yes | | Yes | Yes |
| Employment HHI | | | | | | Yes | | |
| LIML | | | | | | | | Yes |
| Joint F-statistic | 19.0 | 17.9 | 17.7 | 12.9 | 13.4 | 11.5 | 4.6 | 4.6 |
| J-test, p-value | 0.12 | 0.12 | 0.21 | 0.72 | 0.11 | 0.10 | 0.87 | 0.87 |

High-skill concentration (*HS Conc*) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. High-skill concentration threshold $p$ is set at the $24^{th}$ percentile. GDP Growth is real per-capita local GDP growth winsorized at the $1^{st}$ and $99^{th}$ percentiles. Columns (1)-(4) use $B_{i,t-2,large}$ and $B_{i,t-2,small}$ as instruments, each interacted with $\mathbb{1}\{HSConc_{i,t-1} > p\}$. Columns (5)-(7) add to the IV set a SSIV calculated using both small and large shocks together, and using total employment shares as exposure shares. All specifications control for the lagged local sum of shares $s_{i,t-2}$ and are weighted by the twice lagged log of local population. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. Informality refers to the 2000 ratio of informal workers over total employment interacted with year fixed-effects. Non-High-Skill refers to the total non-high-skill hiring instrumented with the SSIVs. Capital Proxy refers to the capital proxy variable calculated using the change in local electricity consumption instrumented with the SSIVs. J-test refers to the overidentification test in Hansen (1982). Municipality-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table A.18: Effect of high-skill concentration in large firms on local skill premium in places with high and low concentration (weighted by log of population)

| | Skill Premium | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| $\mathbb{1}_{\{HS\ Conc._{\cdot t-1}>p\}}$=0 × $HS\ Conc._{\cdot t-1}$ | 9.767** | 10.08** | 8.867** | 10.27** | 10.11** | 10.17** | 9.611** |
| | (3.123) | (3.142) | (3.097) | (3.382) | (3.142) | (3.198) | (3.272) |
| $\mathbb{1}_{\{HS\ Conc._{\cdot t-1}>p\}}$=1 × $HS\ Conc._{\cdot t-1}$ | -1.432** | -1.502** | -1.395** | -1.546** | -1.536** | -1.488** | -1.415** |
| | (0.477) | (0.484) | (0.486) | (0.555) | (0.488) | (0.490) | (0.549) |
| N | 68,582 | 68,582 | 68,582 | 68,582 | 68,582 | 68,582 | 68,582 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | | Yes | Yes | Yes | Yes | Yes | Yes |
| Informality | | | Yes | | | | |
| Non-High-Skill | | | | Yes | | | Yes |
| Capital Proxy | | | | | Yes | | Yes |
| Employment HHI | | | | | | Yes | |
| LIML | | | | | | | |
| Joint F-statistic | 35.9 | 35.5 | 30.9 | 11.1 | 18.9 | 26.4 | 5.0 |
| J-test, p-value | 0.75 | 0.82 | 0.09 | . | 0.57 | 0.70 | . |

High-skill concentration (*HS Conc*) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. High-skill concentration threshold $p$ is set at the $10^{th}$ percentile. Skill premium is defined as the ratio between the average high-skill wage over the average non-high-skill wage within a municipality and it is winsorized at the $1^{st}$ and $99^{th}$ percentiles. Columns (1)-(4) use $B_{i,t-2,large}$ and $B_{i,t-2,small}$ as instruments, the former interacted with $\mathbb{1}\{HSConc_{i,t-1} > p\}$ and the latter with $\mathbb{1}\{\{HSConc_{i,t-1} > p\} = 0\}$. Columns (5) and (6) add to the IV set a SSIV calculated using both small and large shocks together, and using total employment shares as exposure shares. All specifications control for the lagged local sum of shares $s_{i,t-2}$ and are weighted by the twice lagged log of local population. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. Informality refers to the 2000 ratio of informal workers over total employment interacted with year fixed-effects. Non-High-Skill refers to the total non-high-skill hiring instrumented with the SSIVs. Capital Proxy refers to the capital proxy variable calculated using the change in local electricity consumption instrumented with the SSIVs. J-test refers to the overidentification test in Hansen (1982). Municipality-clustered standard errors are in parentheses. $^{*}$, $^{**}$, $^{***}$ denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table A.19: Effect of high-skill concentration in large firms on local GDP growth in places with high and low concentration (twice lagged shares)

| | GDP Growth | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=0 $\times$ $HS\ Conc._{t-1}$ | 1.204** | 1.105** | 1.218** | 1.565*** | 1.055** | 1.056** | 1.601** | 1.631** |
| | (0.375) | (0.371) | (0.394) | (0.459) | (0.363) | (0.366) | (0.499) | (0.512) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}$=1 $\times$ $HS\ Conc._{t-1}$ | -0.462** | -0.448** | -0.477** | -0.586** | -0.429** | -0.421** | -0.600** | -0.611** |
| | (0.157) | (0.154) | (0.163) | (0.181) | (0.154) | (0.156) | (0.192) | (0.197) |
| N | 73,864 | 73,864 | 73,864 | 73,864 | 73,864 | 73,864 | 73,864 | 73,864 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Informality | | | Yes | | | | | |
| Non-High-Skill | | | | Yes | | | Yes | Yes |
| Capital Proxy | | | | | Yes | | Yes | Yes |
| Employment HHI | | | | | | Yes | | |
| LIML | | | | | | | | Yes |
| Joint F-statistic | 14.2 | 14.3 | 12.7 | 12.7 | 13.9 | 13.5 | 5.3 | 5.3 |
| J-test, p-value | 0.10 | 0.09 | 0.17 | 0.52 | 0.12 | 0.12 | 0.54 | 0.54 |

High-skill concentration ($HS\ Conc$) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. High-skill concentration threshold $p$ is set at the $16^{th}$ percentile. GDP Growth is real per-capita local GDP growth winsorized at the $1^{st}$ and $99^{th}$ percentiles. Columns (1)-(4) use $B_{i,t-2,large}$ and $B_{i,t-2,small}$ as instruments, each interacted with $\mathbb{1}\{HSConc_{i,t-1} > p\}$. Columns (5)-(7) add to the IV set a SSIV calculated using both small and large shocks together, and using total employment shares as exposure shares. All specifications control for the lagged local sum of shares $s_{i,t-2}$. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. Informality refers to the 2000 ratio of informal workers over total employment interacted with year fixed-effects. Non-High-Skill refers to the total non-high-skill hiring instrumented with the SSIVs. Capital Proxy refers to the capital proxy variable calculated using the change in local electricity consumption for the service sector instrumented with the SSIVs. J-test refers to the overidentification test in Hansen (1982). Municipality-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table A.20: Effect of high-skill concentration in large firms on local skill premium in places with high and low concentration (twice lagged shares)

| | Skill Premium | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}<p_1\}}=1 \times HS\ Conc._{t-1}$ | 8.088* | 8.660** | 6.955* | 11.49** | 8.582** | 7.126* | 12.09** |
| | (3.307) | (3.357) | (3.311) | (4.037) | (3.319) | (3.128) | (4.606) |
| $\mathbb{1}_{\{p_1<HS\ Conc._{t-1}<p_2\}}=1 \times HS\ Conc._{t-1}$ | -1.429* | -1.498* | -1.247 | -2.054** | -1.492* | -1.136 | -2.172** |
| | (0.669) | (0.682) | (0.673) | (0.780) | (0.681) | (0.612) | (0.779) |
| N | 68,363 | 68,363 | 68,363 | 68,363 | 68,363 | 68,363 | 68,363 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | | Yes | Yes | Yes | Yes | Yes | Yes |
| Informality | | | Yes | | | | |
| Non-High-Skill | | | | Yes | | | Yes |
| Capital Proxy | | | | | Yes | | Yes |
| Employment HHI | | | | | | Yes | |
| LIML | | | | | | | |
| Joint F-statistic | 29.1 | 28.9 | 25.6 | 6.1 | 16.7 | 22.3 | 2.6 |
| J-test, p-value | 0.10 | 0.20 | 0.19 | 0.82 | 0.19 | 0.00 | . |

High-skill concentration (*HS Conc*) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. High-skill concentration thresholds $p_1$ and $p_2$ are set at the $10^{th}$ and $75^{th}$ percentiles, respectively. Skill premium is defined as the ratio between the average high-skill wage over the average non-high-skill wage within a municipality and it is winsorized at the $1^{st}$ and $99^{th}$ percentiles. Columns (1)-(3) use $B_{i,t-2,large}$ and $B_{i,t-2,small}$ as instruments, both interacted with $(\mathbb{1}\{HSConc_{i,t-1} < p_1\} = 1)$ and the large one also with $(\mathbb{1}\{p_1 < HSConc_{i,t-1} < p_2\} = 1)$. Column (4) adds to the IV set a SSIV calculated using both small and large shocks together, and using total high-skill employment shares as exposure shares. Columns (5) and (6) add to the IV set a SSIV calculated using both small and large shocks together, and using total employment shares as exposure shares. All specifications control for the lagged local sum of shares $s_{i,t-2}$. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. Informality refers to the 2000 ratio of informal workers over total employment interacted with year fixed-effects. Non-High-Skill refers to the total non-high-skill hiring instrumented with the SSIVs. Capital Proxy refers to the capital proxy variable calculated using the change in local electricity consumption for the service sector instrumented with the SSIVs. J-test refers to the overidentification test in Hansen (1982). Municipality-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table A.21: Effect of high-skill concentration in large firms on local growth in places with high and low concentration for different thresholds $p$

| | GDP Growth | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=0 \times HS\ Conc._{t-1}$ | 1.091*** | 0.739** | 0.723** | 0.630*** | 0.505** | 0.366** | 0.300* | 0.374* |
| | (0.328) | (0.259) | (0.231) | (0.187) | (0.160) | (0.130) | (0.148) | (0.156) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=1 \times HS\ Conc._{t-1}$ | -0.488*** | -0.480*** | -0.546** | -0.368 | -0.198 | -0.104 | -0.180 | -0.231 |
| | (0.110) | (0.136) | (0.185) | (0.228) | (0.291) | (0.383) | (0.572) | (0.773) |
| Threshold $p$ | 15% | 20% | 25% | 30% | 35% | 40% | 45% | 50% |
| N | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 | 74,090 |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Joint F-statistic | 30.1 | 25.0 | 16.2 | 13.2 | 11.4 | 8.3 | 4.9 | 3.7 |
| J-test, p-value | 0.05 | 0.25 | 0.32 | 0.39 | 0.44 | 0.56 | 0.37 | 0.54 |

High-skill concentration (*HS Conc*) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. GDP Growth is real per-capita local GDP growth winsorized at the $1^{st}$ and $99^{th}$ percentiles. All columns use $B_{i,t-2,large}$ and $B_{i,t-2,small}$ as instruments, each interacted with $\mathbb{1}\{HSConc_{i,t-1} > p\}$. All specifications control for the lagged local sum of shares $s_{i,t-2}$. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. J-test refers to the overidentification test in Hansen (1982). Municipality-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table A.22: Model estimation and moment fit (without the non-innovative, outside firm)

| Parameter | Value | Parameter | Value |
|---|---|---|---|
| $\gamma$ | 1.045 | $\kappa$ | 61268.4 |
| $b$ | 0.62 | $h_l$ | 1.52 |
| $\rho$ | 2092.7 | $h_c$ | 0.32 |
| $A_l$ | 5.97 | | |
| $A_\lambda$ | 61.6 | | |

| Moments | Data | Model |
|---|---|---|
| Growth Rate (%) | 1.31 | 1.01 |
| Skill Premium, Large Firms | 2.76 | 2.77 |
| Labor Market Tightness | 0.48 | 0.00 |
| High-Skill Wage, Non-Large Firms | 0.58 | 0.33 |
| High-Skill Concentration | 0.59 | 0.64 |
| Firm Profitability | 0.20 | 0.25 |
| R&D Investing-to-Sales Ratio (%) | 0.19 | 0.11 |
| Cost-per-Hire | 0.045 | 0.00 |
| High-Skill Unemployment | 0.19 | 1.00 |
| Share of High-Skill Concentration > 80% | 0.38 | 0.32 |

Table A.23: Model estimation and moment fit (1999-2004 period)

| Parameter | Value | Parameter | Value |
|---|---|---|---|
| $\gamma$ | 1.047 | $\kappa$ | 0.44 |
| $b$ | 0.64 | $h_l$ | 1.05 |
| $\rho$ | 5153.5 | $h_c$ | 0.32 |
| $A_l$ | 2.22 | $\nu$ | 0.20 |
| $A_\lambda$ | 36.3 | | |
| **Moments** | | **Data** | **Model** |
| Growth Rate (%) | | 1.30 | 1.29 |
| Skill Premium, Large Firms | | 2.94 | 2.95 |
| Labor Market Tightness | | 0.48 | 0.48 |
| High-Skill Wage, Non-Large Firms | | 0.62 | 0.57 |
| High-Skill Concentration | | 0.45 | 0.53 |
| Firm Profitability | | 0.20 | 0.20 |
| R&D Investing-to-Sales Ratio (%) | | 0.19 | 0.19 |
| Cost-per-Hire | | 0.045 | 0.036 |
| High-Skill Unemployment | | 0.19 | 0.22 |
| Share of High-Skill Concentration $\leq$ 50% | | 0.47 | 0.51 |

Table A.24: Effect of high-skill concentration in large firms on low-skill wages and labor supply

| | Non-High-Skill Wage | | Log(# Non-High-Skill) | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=0 \times HS\ Conc._{t-1}$ | 16.54 | 9.225 | 3.791* | 5.202*** |
| | (677.848) | (690.900) | (1.528) | (1.560) |
| $\mathbb{1}_{\{HS\ Conc._{t-1}>p\}}=1 \times HS\ Conc._{t-1}$ | -8.525 | 21.18 | -0.201 | -0.344 |
| | (121.003) | (126.195) | (0.269) | (0.291) |
| N | 68,607 | 68,607 | 68,607 | 68,607 |
| Time FE | Yes | Yes | Yes | Yes |
| Location FE | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes |
| Informality | | Yes | | Yes |
| Joint F-statistic | 40.3 | 35.6 | 40.3 | 35.6 |
| J-test, p-value | 0.70 | 0.09 | 0.93 | 0.13 |

High-skill concentration ($HS\ Conc$) is the local share of high-skill people working at large firms over total local supply. High-skill workers are those with at least some college education, though they might not have finished their degree. High-skill concentration threshold $p$ is set at the $10^{th}$ percentile. All columns use $B_{i,t-2,large}$ and $B_{i,t-2,small}$ as instruments, the former interacted with $\mathbb{1}\{HSConc_{i,t-1} > p\}$ and the latter with $\mathbb{1}\{\{HSConc_{i,t-1} > p\} = 0\}$. Local-level controls (all lagged to be contemporaneous with instruments): log of population, log of average real wage, the percentage of high-skill workers in the population, the population share of workers receiving minimum wage or less, and the ratio of net hiring over population. Informality refers to the 2000 ratio of informal workers over total employment interacted with year fixed-effects. J-test refers to the overidentification test in Hansen (1982). Municipality-clustered standard errors are in parentheses. *, **, *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

## A.2 Romer-based Growth

In this section, I derive the Romer-model-based growth rate from an increase in human capital supply. Since the goal here is not a full derivation of the model, I stick to the aspects that matter to us, primarily the R&D production function. I will use a similar notation to Jones (1995) which generalizes the original model. Importantly, this means that the parameters here are not related to the ones used in the main text of this paper.

In Romer (1990), the production function of the final good takes the form:

$$Y = L_Y^{1-\alpha} \sum_{j=1}^{A} x_j^{\alpha} \tag{34}$$

where $Y$ is output, $L_Y$ is labor used in the production of the final good, $x_j$ are intermediate goods, $\alpha$ is a constant, and $A$ is the number of intermediate goods. The latter can be thought of as the number of product ideas in the economy.

The total number of intermediate goods evolves according to the following "ideas production function":

$$\dot{A} = \gamma L_A^{\lambda} A^{\phi} \tag{35}$$

where $\gamma$, $\lambda$ and $\phi$ are constants, and $L_A$ is the number of workers engaged in innovation. The labor market clearing condition is, then:

$$L = L_A + L_Y \tag{36}$$

where $L$ is total labor supply.

In steady state, the share of labor employed in R&D is constant. Let $L_A = s_A L$ where $s_A$ is said share. From Equation 35 we can write the growth rate of product variety in this model as:

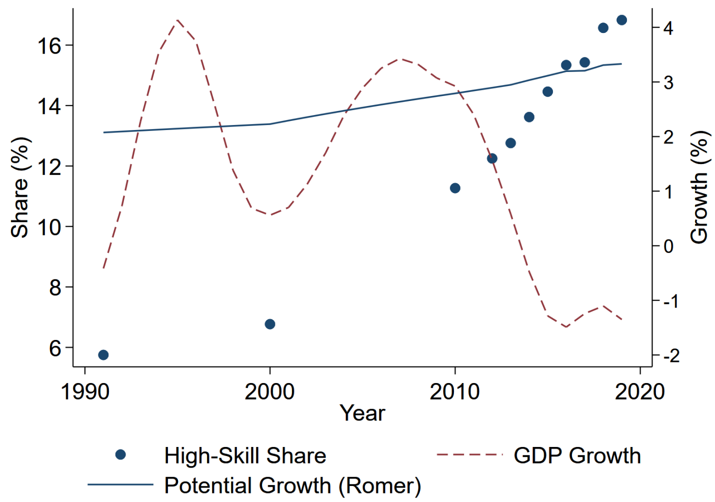
$$g_{Romer} = \frac{\dot{A}}{A} = \gamma (s_a L)^{\lambda} A^{\phi - 1} \tag{37}$$

In the original model, this growth rate is constant implying that the time derivative of the right-hand side of Equation 37 is zero. That is:

$$\frac{\dot{L}}{L} + (\phi - 1)\frac{\dot{A}}{A} = 0 \tag{38}$$

In our case, we assume that $s_A$ grows constantly for a time period as the economy moves towards a new steady state where the labor share employed in innovation is $s'_A$,

Figure A45: *High-skill population share, GDP per-capita growth trend, and Romer-model-based expected growth in Brazil between 1991 and 2019*




Note: High-skill share data is from the Atlas of Human Development in Brazil (UNDP et al., 2024). High-skill share corresponds to the ratio between the number of people with a college degree and the total population who is at least 25 years old. GDP growth is the real GDP per-capita growth trend after filtering data since 1970 using a Hodrick-Prescott filter. Expected Growth is the expected growth from the increase in high-skill labor share in a Romer-based model (c.f. Section A.2).

$s'_A > s_A$. During this transition, we can write the change in the growth rate as:

$$\frac{\dot{g}_{Romer}}{g_{Romer}} = \lambda\left(\frac{\dot{s}_A}{s_A} + \frac{\dot{L}}{L}\right) + (\phi - 1)\frac{\dot{A}}{A} \tag{39}$$

where we now take into account that the share $s_A$ is changing.

If we make the assumption that, throughout the transition, the economy moves between steady states, we can substitute Equation 38 into Equation 39 to get that:

$$\frac{\dot{g}_{Romer}}{g_{Romer}} = \lambda\left(\frac{\dot{s}_A}{s_A}\right) \tag{40}$$

Hence, the change in the growth rate is only due to changes in the share of high-skill labor. We can then use Equation 40 to calculate the Romer-model-based expected growth rate from a change in high-skill supply. This is done in Figure A45 using $\lambda = 0.435$ (Pessoa, 2005).

One important assumption from this calculation is that the only change relative to the original model is the increase in the share of high-skill labor. Naturally, this increase can be countered by changes to parameters $\gamma$, $\lambda$, and $\phi$. For example, a reduction in innovation productivity (i.e. lower $\gamma$) or an increase in the concavity of high-skill labor (i.e. lower $\lambda$) could lead to a constant growth rate. As such, the values in Figure A45 should be interpreted as a measure of *potential growth rate*.

## A.3 Monte Carlo Simulations on Identification with Endogenous Controls

In the baseline estimation in Section 3, we control for non-high-skill hiring (and a proxy for capital formation) to make sure that the exclusion restriction on the SSIV holds, i.e. that any low-skill hiring induced by the SSIV has no significant effect on local GDP growth. This, in turn, might introduce bias if non-high-skill hiring is happening as a result of additional high-skill hiring. This is known as a "bad control" problem. In this section, I show that we can still identify the coefficients of interest when we add endogenous controls and instrument them with the extra instruments.

We start by defining the data generating process. There is a set of four IVs ($Z_{1-4}$), three endogenous variables ($H_1$, $H_2$, and $L_1$), and a dependent variable Y, all of which relate as follows:

$$\begin{aligned}
H_1 &= \beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3 + \beta_4 Z_4 + v \\
H_2 &= \gamma_0 + \gamma_1 Z_1 + \gamma_2 Z_2 + \gamma_3 Z_3 + \gamma_4 Z_4 + w \\
L_1 &= \alpha_0 + \alpha_1 Z_1 + \alpha_2 Z_2 + \alpha_3 Z_3 + \alpha_4 Z_4 + u \\
L_2 &= \delta_0 + \delta_1 H_1 + \delta_2 H_2 + \nu \\
Y &= \xi_0 + \xi_1 L_1 + \xi_2 H_1 + \xi_3 H_2 + +\xi_4 L_2 + \epsilon
\end{aligned} \tag{41}$$

where $L_2$ represents the change in non-high-skill hiring due to changes in high-skill hiring and $(u, v, w, \nu, \epsilon)$ are error terms. Equation 41 can be understood as follows. Instruments $Z_{1-4}$ generate variation in both low-skill hiring ($L_1$) and high-skill concentration, the latter split between low ($H_1$) and high ($H_2$) high-skill concentration places. Low-skill hiring can also happen due to substitutability or complementarity with high-skill hiring ($L_2$). Finally, both low and high-skill workers contribute to output ($Y$). We are interested in identifying $\xi_2$ and $\xi_3$.

We then match all parameter moments to their estimated values in Section 3 and we draw 100,000 joint observations of the IV and error term sets matching their empirical distributions, in particular their in-sample covariance structure. Finally, we estimate $\xi_2$ and $\xi_3$ using 2SLS in three different scenarios: one where we fix $\delta_{0-2}$ and we vary $\xi_4/\xi_1$, one where we set $\xi_1 = \xi_4$ and we vary $\delta_1$, and finally one where vary both $\delta_1$ and the $\xi_4/\xi_1$ ratio. The idea is to assess identification as we vary the intensity of the "bad control" channel with respect to both the effect of high-skill concentration on low-skill hiring ($\delta_{0-2}$) and the effect of the change in low-skill hiring on growth ($\xi_1$ and $\xi_4$). If identification fails, it would be important to determine for which range of parameter values of the unobserved "bad control" channel it happens. Ideally, results will show that we can identify the parameters of interest for any realistic range of $\xi_4/\xi_1$ and $\delta_1$, i.e. the intensity of the "bad control" channel.

I report results in Figure A46 for $\hat{\xi}_2$.[62] On the top-left plot, we fix the intensity of the effect of high-skill concentration on low-skill hiring though we increase the effect of the additional low-skill hiring on growth. On the top-right one, we set $\xi_1 = \xi_4$ and we increase the effect of high-skill concentration on low-skill hiring. Finally, on the bottom plot we increase both channels simultaneously. Importantly, there are two takeaways from this exercise. First, identification starts to weaken as the "bad control" channel becomes more significant in magnitude. We see this in the increasing distance between the horizontal lines representing the parameter's true value and the point estimates. However, unless both the change in low-skill hiring due to high-skill concentration and its effect on growth relative to other low-skill workers are quite large the parameter of interest is identifiable. A value above 8 for both $\delta_1$ and $\xi_4/\xi_1$ would imply an implausibly large high-to-low skill elasticity and that low-skill workers hired through this channel are more productive than other low-skill workers by almost an order of magnitude.

As such, the approach taken in Section 3 of controlling for non-high-skill hiring (and the proxy for capital, as results are analogous) does not seem to introduce a significant bias due to a "bad control" problem under reasonable values for its intensity while allowing us to identify the effect of high-skill concentration on growth.

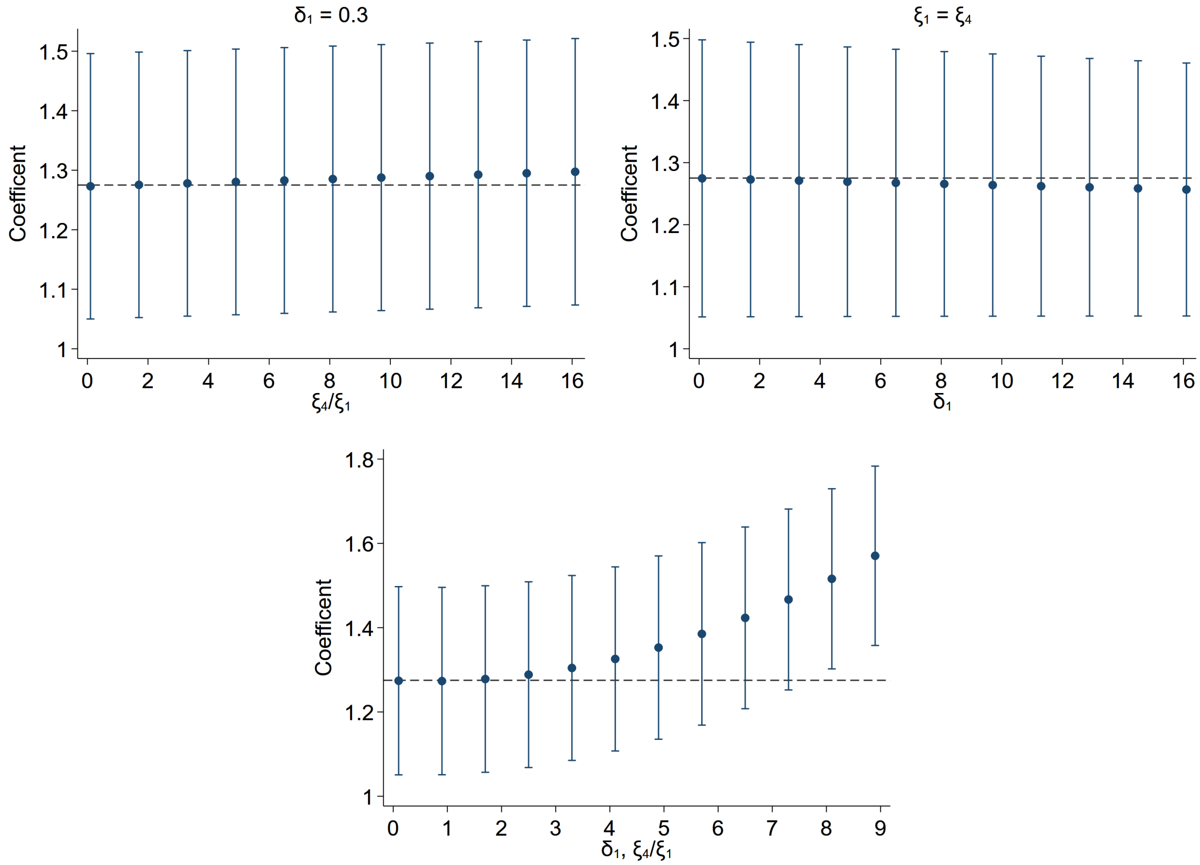
## A.4 SSIV Robustness Checks

We now assess the robustness of our non-monotonic results with respect to changes in the baseline specification. Importantly, what interests us is not exactly the stability of point-estimates but whether the coefficient signs are robust to changes in the estimation, i.e. whether our non-monotonicity result is robust.

We first assess whether having instruments interact with $\mathbb{1}\{HSConc_{i,t-1} > p\}$ leads to biased estimates. As $\mathbb{1}\{HSConc_{i,t-1} > p\}$ is a function of the endogenous variable, albeit one with little variation, we may worry that we might be reintroducing endogenous variation into the instrument set. I report in Table A.9 in the Appendix results where I use polynomial terms for both high-skill concentration and the SSIVs instead of interacting them with $\mathbb{1}\{HSConc_{i,t-1} > p\}$. Although instrument relevance is significantly lower, we observe a similar non-monotonic (and concave) relationship between high-skill concentration and growth. I also report the point-estimates of the point where the slope

[62]Results are similar for the channel through $\hat{\xi}_3$.

Figure A46: *Estimated $\hat{\xi}_2$ for different "bad control" conditions*




Note: Horizontal line corresponds to the true value of $\xi_2$. Vertical bars correspond to the 95% confidence interval.

changes sign to confirm that they are within the [0,1] domain. As a final check, I report in Table A.10 results from a specification that interacts high-skill concentration with $\mathbb{1}\{HSConc_{i,t-1} > p\}$ but instruments with polynomial terms of the SSIV. Results remain robust though we lose significance of the coefficient at high levels of high-skill concentration when the instrument gets weak. Evidence, then, suggests that results are not affected by having instruments interact with $\mathbb{1}\{HSConc_{i,t-1} > p\}$.

Next, we analyze whether results remain unchanged when we use a narrower definition of high-skill. Up to this point, we have considered all workers with some college education as high-skill employees by assuming that, while not all of them are scientists and engineers, they can contribute to incremental, process, and/or catch-up innovation which usually require the type of critical thinking fostered at university. However, we can narrow down this definition to include only those who actually work in occupations that require high critical thinking. Although this procedure may exclude workers who can potentially do innovation yet are underemployed relative to their capabilities, this narrower classification would reinforce the link between our results on high-skill concentration and growth to innovation dynamics. I show, then, in Tables A.11 and A.12 results for the baseline estimation on growth and skill premium, respectively, using high critical thinking workers as defined in Section 2. We find similar non-monotonic relationships between high-skill concentration, growth, and skill premium as in the baseline case.

On skill premium, I show evidence that the estimated negative slope is due to a large negative coefficient at an intermediate level of high-skill concentration. As discussed in Section 3.3, the estimation using skill premium suffers from a weak IV problem with respect to the SSIV for small firms when concentration is high. To assess this issue, I run in Table A.13 in the Appendix a specification where instead of splitting the regressor between low and high levels of high-skill concentration thresholds ($\mathbb{1}\{HSConc_{i,t-1} > p\}$ and $\mathbb{1}\{HSConc_{i,t-1} \leq p\}$, respectively) I do it between low and mid-level concentration ($\mathbb{1}\{HSConc_{i,t-1} < p_1\}$ and $\mathbb{1}\{p_1 < HSConc_{i,t-1} < p_2\}$, respectively). I then fix $p_1$ at the $10^{th}$ percentile of the high-skill concentration distribution and vary $p_2$ for different estimations. Results show that the negative slope at high levels of concentration is due to a more negative coefficient at mid-levels (between -5.0 and -2.5 vs. -1.4 in the baseline). This explains why the overidentification test fails: as we increase the $p_2$ threshold the SSIV for small firms becomes weaker, which can be seen in the joint F-statistic jumping from 38.3 to 55.1 between Columns 4 and 5 once we remove $B_{i,t-2,small}(\mathbb{1}\{p_1 < HSConc_{i,t-1} < p_2\} = 1)$ from the set of instruments, and instruments are possibly capturing the heterogeneity in the slope. Importantly, the failure of the overidentication test does not seem to be due to a violation of instrument validity. I show further evidence of this in Table A.14 in the Appendix where I use $B_{i,t-2,small}$ as an instrument without interaction terms. The non-monotonic result is robust and we cannot reject the null hypothesis of the overidentification test.

I also show that we obtain similar results when we run the specification on a subsample restricted to the non-tradable sector.[63] This is shown in Tables A.15 and A.16 for GDP growth and skill premium, respectively. The fact that we observe similar results for non-tradables is reassuring as tradable firms can engage in product competition with companies outside their municipality. As I show in Section 4, the mechanism I propose to explain the reduced-form results involves the strategic interaction between a leading and a follower firm competing through innovation. Since I run my baseline specification at the municipality level, I also capture tradable firms competing out-of-municipality. By

[63] I define the non-tradable sector as any sector outside agriculture and manufacturing. Although some service subsectors can be deemed tradable, there is no local-level GDP data by subsector.

restricting the analysis to the non-tradable sector, particularly in a context where the data is at the establishment level, results can be more directly linked to my mechanism.

Finally, results are robust to other changes in the specification. We still observe non-monotonicity and significant results between high-skill concentration, growth, and skill premium if we run a weighted regression weighting by the twice lagged log of local population (Tables A.17 and A.18). Results are also robust to lagging the SSIV exposure shares one additional period, i.e. using $s_{in,t-4}$ instead of $s_{in,t-3}$ (Tables A.19 and A.20, the latter uses the low and mid-level thresholds to increase IV relevance). Finally, I show in Table A.21 that while point-estimates are sensitive to the choice of threshold $p$ between lower and higher skill concentration places, our finding on the non-monotonic relationship does not depend on a particular value of $p$ as long as the cut-off is near the point where the slope changes in the relationship between high-skill concentration and growth (or skill premium). In our case, results show that the change in slope occurs around the range of 15% and 25%.

## A.5 Low-Skill Labor Supply Assumption, Proofs, and Derivations

In this section, I discuss the assumption of perfect elastic low-skill labor supply and provide the necessary proofs and derivations for the model derived in Section 4.

*Perfect elastic low-skill labor supply:* As in Aghion et al. (2001), I assume in the model that low-skill labor supply is perfectly elastic. Along with the mathematical simplification, this assumption, in fixing low-skill wage, leads to a straightforward link between variations in the skill premium and what is happening in the high-skill labor market, particularly with respect to demand for high-skill labor. As such, the non-monotonic result on skill premium is only being driven by changes in high-skill wages.

I assess this assumption empirically in Table A.24 using the 2SLS specification in Section 3.3. We see that changes in high-skill labor concentration do not lead to significant changes in low-skill wages. Importantly, the high standard errors in Columns (1) and (2) are due to the low in-sample variance of real low-skill wages across municipality-year pairs. Moreover, Columns (3) and (4) show a positive relationship between high-skill concentration and low-skill hiring when the former is low. These results are in-line with assuming that low-skill wages are fixed while low-skill labor supply adjusts.

In reality, low-skill labor supply is elastic though not infinitely so. Lobel (2024) estimate a 4.15 labor supply elasticity in Brazil which is in-line with assuming perfect elasticity (vs. infinitely inelastic supply). If we had instead assumed a finite labor elasticity, some of the conclusions from the infinitely inelastic case studied in Aghion, Harris and Vickers (1997) would apply. We can expect low-skill wages to follow changes in aggregate demand, which in turn are related to profits, high-skill hiring by the non-innovative sector, and low-skill hiring. Demand, however, varies little across different gap levels as lower low-skill hiring due to higher relative productivity is compensated by higher profits. This implies that the non-monotonic shape of the skill premium curve as a function of high-skill concentration would not change significantly. I assess this point further in Section A.6 by running the model under different labor market assumptions.

*Optimal R&D investment and high-skill labor demand:* The firm's optimal R&D choice can be derived from the maximization problem in Equation 19. Starting with $\lambda_s$, the first-order condition for the leader problem results in (analogously for $\lambda_{-s}$ and $\lambda_0$):

$$\lambda_s = \frac{A_\lambda(J_{s+1} - J_s)}{\rho} \tag{42}$$

Similarly for labor demand $l_{s,H}$, which requires solving:

$$w_{s,H} = A_l \alpha l_{s,H}^{\alpha-1}(J_{s+1} - J_s) - \kappa l_{s,H}\left(\frac{\delta}{B\theta_s^{-\varphi} u_s}\right)^2 \tag{43}$$

where we used Equation 22 to replace for $v_s$ as a function of $l_{s,H}$. To solve for labor demand, we substitute for the wage rate using Equation 27.

*High-skill wage:* To get Equation 27, we multiply Equation 25 by $r$ and replace $W_s$ and $U_s$ with their definitions along with the Nash bargaining solution which helps us get rid of $W_s$ and $U_s$:

$$r\xi S_s = rW_s - rU_s = w_{s,H} - \delta(\xi S_s) - b - B\theta_s^{1-\varphi}\xi v_s S_s - B\theta_s^{1-\varphi}\xi v_{-s} S_{-s} \tag{44}$$

where we used the fact that the match surplus for the non-innovative firm is zero. We can then rearrange terms to get Equation 27.

*Growth rate:* The derivation follows Acemoglu and Akcigit (2012). Start with a single sector with a gap $s$. Since $y_s = \gamma_s l_{s,H}$ and $l_{s,H}$ is constant in steady-state, $y_s$ grows at the same rate as $\gamma_s$, i.e.:

$$g_s = \lim_{\Delta t \to 0} \frac{\ln\gamma_s(t+\Delta t) - \ln\gamma_s(t)}{\Delta t} \tag{45}$$

Given Bertrand competition, we only need to look at the leader's production for $s > 1$ and the neck-and-neck case, though we will link it back to the follower's case at the end. Note that at any interval $\Delta t$, in expectation, the leader innovates at a rate $\eta_s \Delta t + o(\Delta t)$ while neck-and-neck firms innovate at a rate $2\eta_0 \Delta t + o(\Delta t)$. Each innovative step increases $\gamma_s$ by $\gamma$. Then:

$$\ln\gamma_s(t+\Delta t) = \ln\gamma_s(t) + \ln\gamma\Big[\mathbb{1}_{s=0}2\eta_0\Delta t + \mathbb{1}_{s>0}\eta_s\Delta t\Big] \tag{46}$$

Replacing Equation 46 into 45 results in Equation 31.

The final step is to notice that aggregate growth is the weighted average of all $g_s$ by the sector share $\mu_s$. Notice also that, in steady-state, the technological frontier (i.e. leaders and neck-and-neck firms) and followers must grow at the same rate, implying that:

$$g = \ln(\gamma)\Big(\sum_{s=1}^{\infty}\mu_s\eta_s + 2\mu_0\eta_0\Big) = \ln(\gamma)\sum_{s=1}^{\infty}\mu_s[\eta_{-s} + h_l l^{\alpha}_{-s,H} + h_c] \tag{47}$$

## A.6 High-Skill Labor Supply Assumption

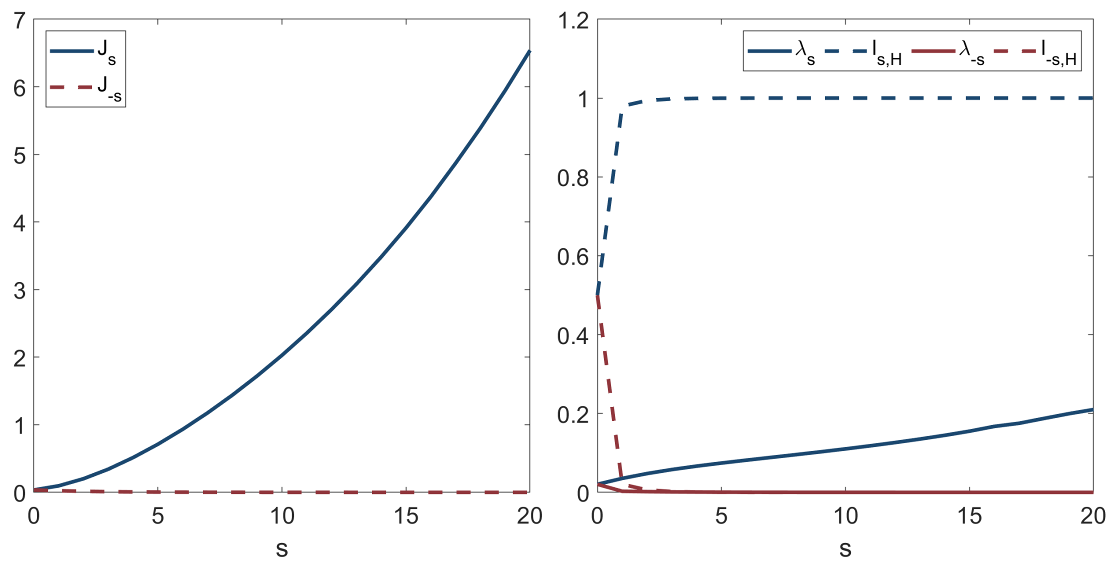
In the model presented in Section 4, high-skill labor is hired through search. In this section, I show how results change if we remove labor frictions, providing intuition for their importance. I then show the importance of labor frictions in matching the skill premium.

We first consider a version of the model with high-skill labor which is closest to Aghion et al. (2001). That is, we start with a similar set-up to the one in Section 4, i.e. a step-by-step growth model with two firms, and we allow for two types of labor, high and low skilled, each being paid at wage $w_k$, $k = \{H, L\}$. As in Section 4, high-skill labor is used in R&D production. Importantly, there are no search frictions, hence no unemployment. This implies that the labor market for high-skill workers has to clear as follows:

$$L_H = l_{s,H} + l_{-s,H} \tag{48}$$

Note how Equation 48 implies that total high-skill hiring ($L_H$) is invariant with respect to the gap $s$. This aspect of this simple model affects results significantly. To see that, first realize that in this model the leading firm still has relatively higher incentives to hire skilled workers than the follower as $s$ increases similar to results in Section 4. This is because we have not made any changes to firm competition or how innovation works. As such, starting at $s = 0$, as $s$ increases high-skill concentration at the leader still goes up. However, once incentives to innovate decline as the leader is too far ahead, the leading firm cannot shed labor as the skilled labor market has to clear and the follower is even less willing than the leader to hire. As such, $l_{s,H}$ is necessarily a monotonically increasing function of the gap $s$. I show this in Figure A47 where I plot the firms' value functions

Figure A47: *Model results without labor search and an outside sector*




Note: $J_s$ ($J_{-s}$) refers to the value function of the leader (follower). $\lambda_s$ ($\lambda_{-s}$) refers to R&D investment by the leader (follower). $l_{s,H}$ ($l_{-s,H}$) refers to high-skill labor hired by the leader (follower).

and input decisions as a function of the gap $s$.[64] We observe that the leading firm' high-skill labor hiring increases monotonically in $s$ and stays near total supply $L_H$ for $s$ large enough. Notice moreover that this change in hiring relative to the model in Section 4 (Figure A27) also affects the investment decision.

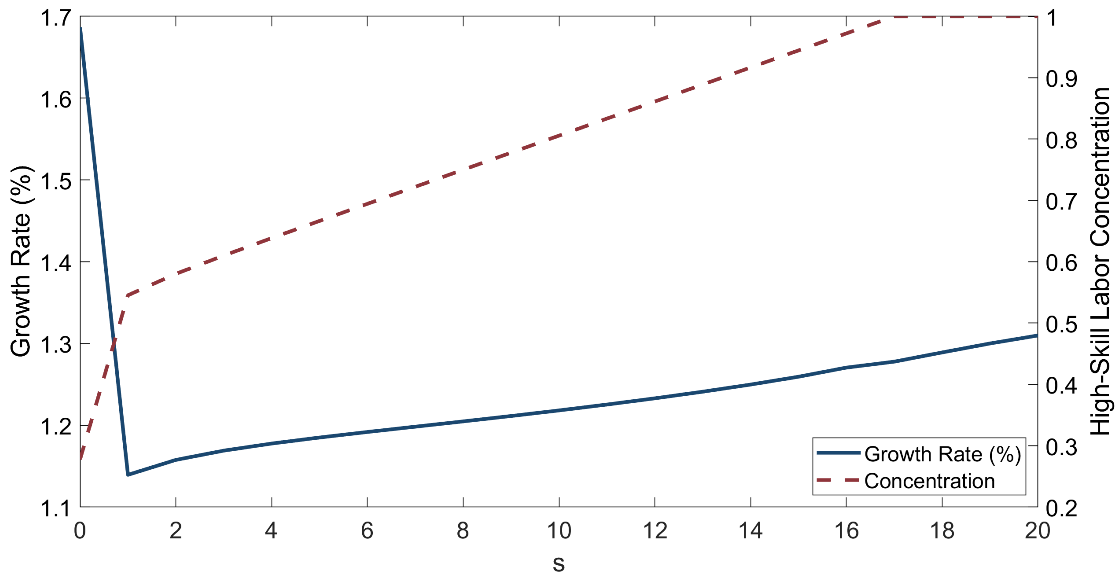
This change in the skilled labor market clearing also affects results on growth. I show this in Figure A48. Since the leading firm cannot lower its high-skill hiring, growth does not decline at a high-enough level of the gap. Key to this is that labor cannot adjust downwards as it does in the baseline model. This highlights the importance of allowing for unemployment in the model.[65]

Labor frictions are not only important for results on growth but also to match the skill premium. To show this clearly, we can take results from Section 4, that is those from a model with high-skill labor search, and counterfactually change the assumption on high-skill labor.

[64]To solve the model, we normalize low-skill labor supply to 1. As it is simpler, this model only has 5 parameters to be estimated: $\{\gamma, \rho, A_l, A_\lambda, h_c\}$. I estimate those using 5 moments: average real GDP per capita growth rate, average high-skill labor concentration at large firms, average firm profitability, R&D share of sales, and share of markets where high-skill concentration is below or equal to 50%.

[65]Results are similar if we go another step forward and add an outside sector as in Section 4. Though this sector can in theory absorb high-skill labor once the leading firm faces lower incentives to hire, this is limited by the non-innovative sector's demand. Results for the version with an outside sector are available upon request.

Figure A48: *Growth and high-skill labor concentration in the model without labor search and an outside sector*



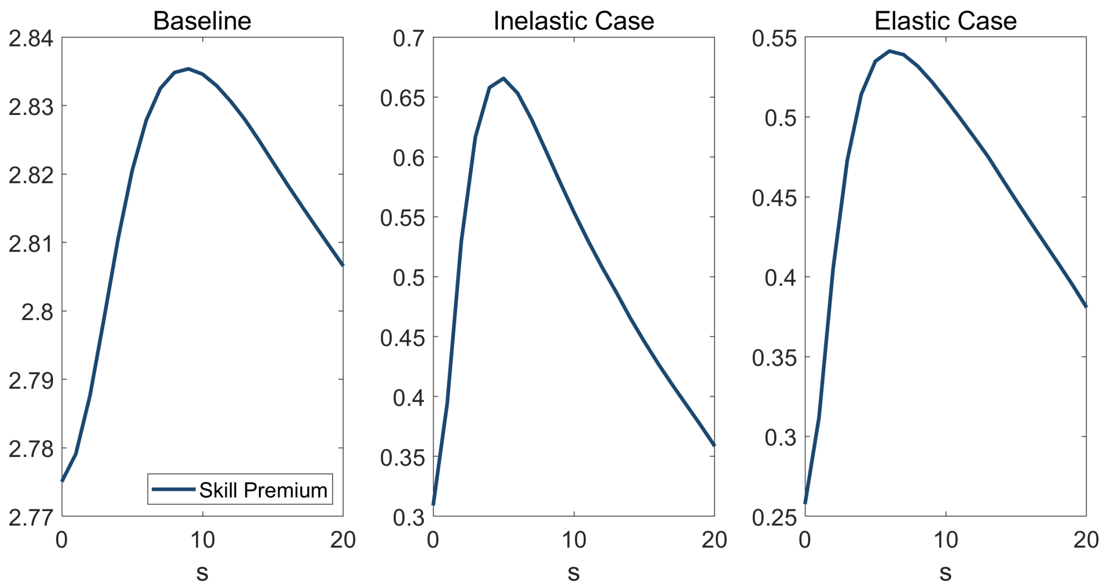

We start with a similar set-up to the one in Section 4, i.e. a step-by-step growth model with two types of labor (high and low skilled). However, we assume now that high-skill labor can be hired without frictions and is supplied inelastically. All firms pay a single wage rate, conditional on the gap $s$, which is set so that the labor market clears. Taking the first-order condition with respect to labor demand in Equations 19 and 23 we get the following result:

$$
\begin{aligned}
l_{s,H} &= \left(\frac{A_l\alpha(J_{s+1} - J_s)}{w_{s,H}}\right)^{\frac{1}{1-\alpha}} \\
l_{-s,H} &= \left(\frac{A_l\alpha(J_{-s+1} - J_{-s})}{w_{s,H}}\right)^{\frac{1}{1-\alpha}} \\
l_{0,H} &= \left(\frac{A_l\alpha(J_1 - J_0)}{w_{0,H}}\right)^{\frac{1}{1-\alpha}} \\
l_{o,H,s} &= (1-\nu)\frac{D_s}{w_{s,H}}
\end{aligned}
\tag{49}
$$

Intuitively, Equation 49 says that high-skill labor is paid at its marginal product which, for the innovative sector, is the marginal benefit, in expectation, from a successful R&D innovation that increases the gap by 1. Importantly, in this case total labor demand does not vary with $s$ as there are no restrictions in the labor market. We can then solve Equation 49 using, where needed, estimates from the baseline estimation, including the estimates for the firms' value functions.[66] I plot results for the wage premium in Figure A49. Com-

[66]To make results comparable, I set total labor supply $L_H$ to average employment in the baseline model.

Figure A49: *Wage premium as a function of the gap s for different assumptions on high-skill labor supply*




paring the wage premium curve between the inelastic and labor search ("baseline") cases, we observe that we achieve non-monotonicity in both cases as firms move from the region of intense competition to the one where the lazy monopolist effect kicks in. This is because we are using the results from the model with search frictions, which allow high-skill labor hiring to adjust. However, even when we use the baseline results for the firms' value function there are two shortcomings of the inelastic case. First, wage premium can be (and is) below one which does not make empirical sense. Second, since at large $s$ both firms have low incentives to invest in R&D effort, the high-skill wage approaches zero.[67] As such, the wage premium also goes to zero as $s$ grows large which also does not reflect reality.

Another approach would be to make high-skill labor supply elastic by adding labor disutility to the utility function of workers. The shape of the wage premium curve in this case would depend on the exact functional form of the utility function. In cases where income and substitution effects cancel out, the end result is a constant $L_H$ and a simple level shift from the inelastic case. However, in cases where labor supply is a monotonic function of the wage the shape of the wage premium curve changes slightly though it is still dictated by the change in the marginal benefit of R&D. I show one parametrization of the latter in Figure A49 where I set the Frisch elasticity $\zeta$ to 0.5 and the disutility scalar is set to match the average unemployment rate in the baseline case (i.e. with high-skill

[67] In the limit, for $s$ reaching infinity high-skill wage is effectively zero as the value function becomes flat.

labor search).[68] In this scenario, as the disutility from working at different companies is the same we have to impose a geographical restriction to the labor market where firms hire from within different areas in a municipality.[69] The resulting wage premium is also below 1 and tends to zero as $s$ grows.

It is clear from Figures A48 and A49, then, that the assumption of search frictions is helpful, both qualitatively and quantitatively. Qualitatively, having unemployment allows firms to adjust high-skill labor downwards, in tandem with changes in incentives to hire as a function of the gap $s$. Quantitatively, the baseline scenario can capture empirical trends, particularly when it comes to an above-one skill premium, and does not rely on a particular shape of the disutility of labor. Moreover, the elastic case still requires an assumption on labor mobility as firms pay different wages, which is an empirical fact. Finally, we also require some restriction in the labor market to capture unemployment. Search frictions are, then, a natural choice.

## A.7 GMM Estimation Moments

In this section, I go over each empirical moment and the model mapping of the GMM estimation in detail:

i) *Real GDP growth*: data comes from IBGE for the 1999-2017 period. I show how to calculate the aggregate growth rate in the model in Section 4;

ii) *Skill Premium at Large Firms*: average calculated using in-sample data where I weight observations by the number of workers. In the model, skill premium is the labor weighted average of high-skill labor in firms $i$, $-i$, and $o$;

iii) *Labor Market Tightness*: data comes from the Catho-Fipe series which provides an indexed time series. The average level between 2004-2017 is calculated using a nominal value reported in October 2013, allowing me to de-index the data;

iv) *High-Skill Wage at Non-Large Firms*: calculated using in-sample data weighted by the number of workers. Wage data is provided as a multiple of the yearly minimum wage. To recover annual wage rates, I multiply the wage multiple by the monthly minimum wage rate from IBGE. I then multiply it by 13 to get annual wages, taking into account the mandatory end-of-the-year bonus. To avoid an empirical moment with a large order of magnitude, I divide the annual wage by 100,000;

v) *High-Skill Concentration*: average calculated using in-sample data. I show how to

[68] An example of a case where $l_{s,H}$ does not depend on the wage is $U(x,L) = \ln(x) - K\frac{l_{s,h}^{1+\frac{1}{\zeta}}}{1+\frac{1}{\zeta}}$ which delivers $l_{s,h} = (\frac{1}{K})^{\frac{\zeta}{1+\zeta}}$. In the case where $l_{s,H}$ depends on the wage, I abstract from the exact functional form and set $l_{s,h} = (\frac{w_{s,H}}{K})^{\zeta}$ where $K = 3.62$.

[69] Otherwise, all labor would work at the leading firm, who pays a higher wage.

calculate high-skill concentration in the model in Section 4;

vi) *Firm Profitability*: calculated as the sum of returns on riskless assets and the equity risk premium (ERP). For the riskless asset, I use $r = 8\%$ as explained in Section 4. For the equity risk premium, I use the value provided in the dataset of Damodaran (2023) (July/23 edition) for Brazil, i.e. $ERP = 9.57\%$. This value, however, is post-tax. Since the model does not take taxation into account, I convert it to the pre-tax level using an effective corporate tax rate of 18.08% (Pires, Marques and Bergamin, 2023). In the model, firm profitability is defined in Equation 14, which must be averaged using the sector shares $\mu_s$;

vii) *R&D Investment-to-Sales Ratio*: data is from the Survey for Technological Innovation (PINTEC) which is conducted over a period of three years since 1998. I use total spending in internal R&D as my measure of R&D investment though it requires two adjustments. First, it is important to take into account government subsidies to R&D spending which account for around 11% of private spending (Betarelli Junior, Faria, Gonçalves Montenegro, Bahia and Gonçalves, 2020). Since there are no subsidies in the model, I remove this share from total private spending. Second, the survey-measured spending includes wages to people employed in R&D activities which should not be taken into account here as we separate investments from labor costs in the model. While the wage share is not measured by the survey, it is known to be substantial as innovation relies heavily on the knowledge capital of high-skill workers. I assume this share to be two-thirds (67%) in line with the literature and US data (Hall and Lerner, 2010, Moris and Shackelford, 2023). In the model, I fit this ratio with the share of total R&D investment over aggregate demand for the innovative sector;

viii) *Cost of Hiring per Job*: I estimate this using data from the US where the average cost of hiring per job was \$4,683 in 2021.[70] I then calculate the cost of hiring as a share of the annual average wage (\$67,610 in 2021 according to the Bureau of Labor Statistics). Finally, I estimate the cost for Brazil as proportionate to the number of days that it takes to hire someone (39.6 days in Brazil vs. 23.8 in the US).[71]. In the model, I calculate this as the average vacancy costs share of high-skill wages;

ix) *Unemployment of High-Skill Workers*: data comes from IBGE for the period between 2012 and 2017 for people who have at least attended college though might not have graduated from it. I adjust average unemployment to take into account that some

[70] Average is from surveys conducted by the Society for Human Resource Management (SHRM).

[71] Data on length of hiring process is from a Glassdoor survey in 2017 available at `https://www.glassdoor.com/research/time-to-hire-in-25-countries`.

high-skill workers are in the informal market, which is something the model does not account for. According to Veloso, de Holanda Barbosa Filho and Peruchetti (2022), around 25% of workers with 16 or more years of study (equivalent to having a college degree) were informal workers between 1999 and 2017. I assume counterfactually that in the absence of an informal market half of currently informal workers would become unemployed (vs. formally employed or becoming inactive). The targeted moment is, then, average high-skill unemployment (6.07%) plus half of those who are informal workers (12.5%);

x) *Share of Markets with Concentration Below 50%*: calculated in-sample after removing markets where concentration is either below 10% or above 90% as those are not properly captured in the model.

Finally, the outside moment ("R&D Worker Share") is calculated using data from PINTEC which reports the number of full-time workers employed in R&D activities and total number of workers for a sample of manufacturing and high-technology firms. In the model, I calculate the R&D worker share as the share of workers employed in R&D at both leader and follower firms.