Why do Firms Train?

Evidence on the Role of Market Power in the Firm's Decision to Train

Ana Catarina Gaspar

Dissertation written under the supervision of Professor Joana Silva

Dissertation submitted in partial fulfilment of requirements for the MSc in Economics, at CATÓLICA-LISBON School of Business & Economics
June 2022

[Page intentionally left blank.]

Why do Firms Train? Evidence on the Role of Market Power in the Firm's Decision to Train

Ana Catarina Gaspar

June 2022

Supervisor: Professor Joana Silva

Abstract

This thesis provides a comprehensive assessment of the relationship between labor market power and firms' decision to engage in job-training. We use rich matched employer-employee data for Brazil and link it to detailed records on training activity from the main provider. To guide our analysis, we start by proposing an extension of a search model to an economy with training, with heterogeneous productivity levels. Our model predicts that larger firms are more likely to train their workers. Constructing a measure of market concentration for all Brazilian labor markets, we find that training is more prevalent in more concentrated markets and that labor market share is an important predictor of training take-up rates at the firm level. Using an instrumental variable approach, we show that training leads to positive average wage gains for workers, but that these gains are smaller for workers employed in high market power firms. Our results shed new light on the role of labor market power in training provision and its impact for workers, with important policy implications for the design of effective labor market programs.

Keywords: on-the-job training, wages, labor market concentration, human capital

[Page intentionally left blank.]

Why do Firms Train? Evidence on the Role of Market Power in the Firm's Decision to Train

Ana Catarina Gaspar

June 2022

Supervisor: Professor Joana Silva

Resumo

Esta tese apresenta uma avaliação abrangente da relação entre o poder no mercado de trabalho das empresas e a promoção de ações de formação profissional. Utilizamos microdados administrativos ligando trabalhadores e empresas para o Brasil cruzando-os com os registos dos trabalhadores a quem foi dada formação profissional pelo principal provedor nacional. Para orientar a análise empírica, começamos por propor uma extensão de um modelo de "search and match" a uma economia com formação profissional, considerando que os níveis de produtividade entre empresas são heterogéneos. O nosso modelo prevê que as empresas maiores têm uma maior probabilidade de dar formação profissional aos seus trabalhadores. Construindo uma medida de concentração de mercado de trabalho para todos os mercados de trabalho brasileiros, mostramos que a formação profissional é mais prevalente em mercados com maior concentração de poder de mercado e que a quota de mercado de uma empresa num certo mercado de trabalho é um importante determinante da probabilidade de promover formação profissional. Utilizando uma variável instrumental, mostramos que a formação profissional conduz a ganhos salariais médios positivos para os trabalhadores, mas que estes ganhos são tanto menores quanto maior o poder de mercado do empregador. Os nossos resultados evidenciam a importância do o poder de mercado para a promoção de formação profissional e o seu impacto para os trabalhadores, com importantes implicações para o desenho de políticas públicas que promovam mercados de trabalho mais eficazes.

Keywords: formação profissional, salários, concentração no mercado de trabalho, capital humano

[Page intentionally left blank.]

Acknowledgements

I would like to start by thanking Professor Joana Silva for all the guidance, support and, especially, for believing in me since the day we met. Besides being a brilliant researcher, Professor Joana is one of the kindest people I have worked with. I also thank Católica and all the outstanding Professors I've had the opportunity to learn from. A special note goes to Professor Pedro Raposo for the feedback on this dissertation. I would also like to mention Professor Isabel Correia and Professor Pedro Teles, who inspired me to continue studying Economics.

A special thanks goes to Jaime Montana, for all his time and extraordinary patience during this process, but also for being such a good friend and making us laugh every day without exception. I also thank Martim Leitão for all the incredible support and for helpful comments and feedback. I am grateful to the great PROSPER team, for all the useful discussions, and for hearing me present this dissertation (way too) many times.

I am thankful to my friends, especially Catarina and António, for ensuring me that I was able to finish this work. A very special thanks to Pedro for his unconditional support and understanding and for listening to me going on about job training and validity of instruments every single day for the past year without ever complaining. Thank you for keeping me grounded, but also for making me dream higher.

Lastly, but most importantly, I thank my parents for always believing in me and for giving me all the tools to become whoever I want to be. All my achievements are yours as well. I dedicate this work to my mother, who has always been my inspiration. A special thanks to my sister who, despite not believing this is what Economists do (in her words, "it seems you are hacking computers with so much coding!"), has always supported me unconditionally.

[Page intentionally left blank.]

Contents

List of Tables

List of Figures

1 Introduction

Is it possible that a firm prefers not to train a worker, even if the productivity gains outweigh the training costs? The answer is yes. Productivity enhancement is only part of the reason why firms train. In fact, what enters the employer's decision to train is how much he/she can *profit* from that increase in productivity. This subtle - but key - difference has been at the center of a growing literature on training that tries to explain asymmetric take-up rates across apparently similar firms (Lynch, 1994; Muehlemann and Wolter, 2011). Labor market competition is at the core of this asymmetry: when firms have market power, they are able to set wages below the marginal factor productivity. The wedge between productivity and wages is what allows firms to recoup some of the training costs, increasing their incentives to invest in it. While this relationship has been acknowledged by the economic theory (Acemoglu, 1997; Acemoglu and Pischke, 1999), the interplay between labor market power and training incentives at the firm level remains relatively unexplored empirically, undermining efficient resource allocation and effective labor market policy design.

This thesis fills this gap by providing a comprehensive assessment of the relationship between market power and firms' decision to engage in job training. Furthermore, we characterize the differential effects of job training on wages depending on market concentration. The literature on training has historically relied on small samples or employee surveys on training programs and is many times limited to cross-sectional identification, due to the lack of rich data on training provision (Cassagneau-Francis et al., 2020; Ibrahim et al., 2020). This, coupled with the heterogeneity of training programs, has compromised the generalization of empirical findings. We use rich matched employer-employer data for brazil (RAIS) and administrative data on one of the largest training programs in the world (SENAI) to build a panel dataset covering all formal firms in Brazil. The Brazilian setting provides a unique opportunity to study the employer's decision to train. First, training programs are designed to meet the demand for skills in an occupation or region, rather than tailored for a specific firm. This general nature of training is particularly relevant for establishing a link between training provision and market power, since all the firms operating in a same market are equally eligible to undertake this investment. Second, a large part of the training cost is both independent of the amount of training provided and paid beforehand in the form of a mandatory contribution. This reduces – and homogenizes – the *per-worker* training costs that firms face and allows us to place a special focus on the most relevant market-specific cost the employer faces when making a training decision: the threat of employee turnover. This underlying threat hinders the firm's ability to recoup its investment in training.

Previous studies have argued that, if skills are transferable across firms, then there is some probability that other firms will poach skilled workers. This may hinder the firm's ability to reap the full benefits of training (Lynch, 1994), leading to under-investment in training. The literature has highlighted that the size of the so-called *"poaching externality"* (Booth and Snower, 1996) is directly linked to labor market density (Muehlemann and Wolter, 2011): In particular, the larger the threat of poaching, the lower the incentives to train. The idea that firms train more when they can extract higher returns is confirmed by studies that find a negative correlation between wage premia and the likelihood to train (Bassanini and Brunello, 2008). There is also evidence that training may be more prevalent in regulated labor markets than in markets in which hiring and firing costs are smaller (Lynch, 1994). Aligned with these insights and to guide our empirical analysis, we start by laying a theoretical model, applying Burdett and Mortensen's wage posting model to an economy with training, based on the work of Manning (2013). We propose an extension of a search model to an economy with training, with heterogeneous productivity levels. In equilibrium, more productive firms can pay higher wages, which attracts, and retains, a larger labor share. Since these larger firms have lower separation rates, investing in training becomes more profitable. Thus, our model predicts that larger firms are more likely to train their workers.

The empirical analysis in this thesis proceeds in two parts. First, we establish a set of empirical facts about the firm-level drivers of training, placing special emphasis on market power. The fineness of our data allows us to define labor markets using occupation and geographical zone, and to study concentration in all Brazilian labor markets. constructing a measure of market concentration for all Brazilian labor markets, we find that training is more prevalent in more concentrated markets and that labor share is an important predictor of training take-up rates at the firm level. Looking closely at the relationship between concentration and training at the firm level, we provide new explanations for different training take-up rates across employers. We find that besides being more frequent in less competitive labor markets, training is significantly more prevalent among firms that have a higher market power, measured by their market share. This is in line with the theoretical prediction that employers train when they believe that separation is less likely given the strong negative correlation between market power and the rate of separations.

Second, we identify the impact of training on the wage distribution focusing on the role of market power. Indeed, we expect that firms who train more are also better able to extract gains from training, perhaps by paying a lower wage premium to trained workers than their similar counterparts with lower market power. The main challenge in estimating the effect of training on key outcomes is the endogeneity of training resulting from self-selection into it. It is known that individuals, and firms, that engage in training differ significantly from those that do not, both in terms of observable and unobservable characteristics. Therefore, we implement an instrumental-variable (IV) approach, instrumenting the firm decision to train with the yearly variation in the availability of training in the relevant labor market. Controlling for time-invariant characteristics of each market and sector-specific unobservable variations over time allows us to meet the exclusion restriction for a valid IV, since the identification comes from intra-market variations in the supply of training across years. Furthermore, we find that changes in training availability are strong predictors of the establishment's decision to train at least one worker, confirming the relevance of our instrument.

Through the instrumental variable approach, we show that training leads to positive average wage gains for workers at the establishment-occupation level, but that these gains are smaller for workers employed in high market power firms. In line with the theory, we also find that training increases wage dispersion inside the firm. The increase in wage inequality is driven by an increase in inequality in the bottom half of the wage distribution. These findings are in line with the literature that finds positive wage returns for workers trained (Bartel, 1995; Blanchflower et al., 1994) and are compatible with the general finding that training is more prevalent when wages are more compressed (Bassanini and Brunello, 2003), since they do not imply that the increase in wage dispersion is higher than the increase in productivity dispersion among workers.[1].

To understand whether the returns to training are different depending on the market power of the firm, we add market share, and its interaction with a measure of training, to the regression estimating training impact on average wages. The resulting estimates provide three important insights. First, establishments with higher market share tend to pay higher average wages. This is aligned with the predictions of our theoretical model that anticipates that, if firms are heterogeneous in productivity levels, then larger firms will pay higher average wages and have a higher market share. Second, the pure effect of training on average wages drops significantly in this regression, which suggests that the previous estimates were biased upwards. This could be partly due to the correlation between training rates and market power. These findings are consistent with our theoretical model, where larger firms pay higher average wages regardless of the amount of training, simply because workers are more productive in these firms. Lastly, we find that the wage returns of training are smaller for workers employed in high market power firms. In fact, among high market power firms, the difference between wages of establishments that train and establishments that do not train is not statistically different from zero.

This thesis contributes to two strands of the literature. First, it contributes to the literature on the drivers and returns to job training. The existing empirical literature in this field focuses mainly on the worker and finds increases in both wages and productivity following employer-provided training programs, (Almeida and Faria, 2014; Bartel, 1995; Bastos et al.,

[1]In this thesis, we use the term wage compression in the sense used in Bassanini and Brunello (2003): There is wage compression if the increase in wage after training is lower than the increase in productivity from it. Therefore, the fact that skilled workers receive a wage higher than unskilled workers does not imply that there is no wage compression.

2016; Blanchflower et al., 1994). At the firm level, there is a consensus that training boosts productivity (Almeida and Carneiro, 2009; Dearden et al., 2006; Konings and Vanormelingen, 2015), but the effect on average wages is smaller, and many times not statistically different from zero (Conti, 2005). Some studies have found a negative relationship between training provision and labor market competition (Lynch, 1994; Muehlemann and Wolter, 2011), but at a more aggregated level. A relatively small number of papers provide evidence on the returns to training using the same dataset used in this thesisr. Particularly, Blyde et al. (2019) focus on the reallocation effect of training courses for unemployed workers, and Bastos et al. (2016) provide a comprehensive analysis of the relationship between export participation and job training. We differ from these studies, however, by shifting our focus to the firm-level determinants and outcomes of training, and studying the impact of the labor market structure.

Secondly, it contributes to the emerging literature on the link between market power and wages. Studies such as Azar et al. (2022), Card et al. (2018), and Webber (2015) have documented that labor market concentration leads to significantly lower earnings for workers. The theory on training predicts that in these markets firms should be able to recoup some of the investment in training by increasing wages by less than the productivity increase resulting from training. In this line, some studies find a negative relationship between wage premia and the likelihood that the firm trains (Bassanini and Brunello, 2008). However, to the best of our knowledge, we are the first to empirically measure the relationship between firm-level market power and training, and to directly examine its consequences on wage distribution. Thus, this thesis complements this literature by unveiling that returns to training are lower the higher the market power of the firm. As far as we are aware, we also provide for the first time a geographical mapping of labor market concentration in Brazil.

Our results shed new light on the role of the strongest player in employee training: the firm. In particular, the findings highlight the idea that promoting on-the-job training requires that policymakers focus on the role of employers and their incentives to engage in such programs. Previous studies have documented that increased competitiveness in labor markets promotes higher wages (Azar et al., 2022) and lower inequality (Leitao et al., 2022). However, our findings point toward a potential unintended consequence of "deregulating" labor markets, since we find that labor market competition reduces incentives to promote general human capital formation.

The remaining of this thesis is organized as follows: Section 2 presents our theoretical framework; Section 3 describes the main datasets used and presents the descriptive statistics. Section 4 explains the methodology adopted to define labor markets, describes the main measures used to characterize them and presents the empirical model. Section 5 presents the empirical model and describes the main drivers of the establishment's decision to train, focusing on the market structure. Section 6 presents the instrumental variable used and proceeds by describing training impacts on average wages and on their dispersion, as well as their relation to market power. Section 7 presents the results of some robustness checks, and Section 8 concludes.

2 Why do firms train their workers?

Becker's foundational work on the theory behind training dominated economic thinking in the field of human capital accumulation for many years. The conclusions and empirical predictions in the paper are based on the simple distinction between specific training (training that increases workers' productivity only with the current employer) and general training (training that increases workers' productivity with more than one employer). Strikingly, the paper predicts that, assuming a perfectly competitive market structure, employers will never provide training simply because they are not able to recoup their investment afterward. Therefore, if it is observed that an employer incurs a cost to providing training then it must either be because training is specific or because workers are indirectly paying for their training by receiving lower wages.

Over the years, the existence of several training programs where firms were paying for training, and skills were transferable continued to defy the existing consensus. It was not until 1999, in the work of Acemoglu and Pischke (1999), that a discussion was opened on the role of the market structure (a key assumption in Becker's model) for training. The reason is simple: If firms have some power over their workers, then the supply of labor directed at each firm is not perfectly elastic, and firms can pay wages lower than marginal productivity, which implies that they are able to collect some of the returns to training. Therefore, the power of the firm in the labor market may be an important explanation for why employer-provided training is observed in many markets, even when skills are perfectly general.

2.1 Theoretical Framework

To motivate our empirical analysis, we start by laying out a theoretical framework that describes firms' decision to train in a monopsonistic market. Our model explores a simple mechanism that generates a correlation between market shares and training rates at the firm level, based on heterogeneity in productivity levels between firms. Following a modeling strategy consolidated in the literature (see Burdett and Mortensen, 1998) and adapted to an economy with training by Manning (2013), we work with an equilibrium search model where employers can provide general training to convert an unskilled worker into a skilled one. Our model differs from the one in Manning (2013) since we generalize it to the case where there is a continuous distribution of productivity levels across firms and assume a specific form for the relationship between the productivity of unskilled and skilled labor. Dispersion in wages in the economy, however, does not depend on any of these assumptions: in Burdett and Mortensen's model, a distribution of wages arises even if firms are symmetric at the outset.

We consider an economy where there is a unitary mass of workers, all of them equal at the outset. All workers are unskilled ($T = 0$) but they can become skilled ($T = 1$) if they are trained by the firm (and this change is permanent). There is a unitary mass of firms, whose technology has constant returns to scale. Firms differ in productivity levels. The production function of firm i is[2]:

$$Y_i = p_{i0}N_{i0} + p_{i1}N_{i1} \tag{1}$$

where p_{i0} is the productivity of unskilled workers, p_{i1} is the productivity of skilled workers, and N_{i0} and N_{i1} denote the amount of unskilled and skilled labor, respectively. There is a continuous distribution of p_{i0} in the economy. We further assume that training increases the productivity of a worker at any firm by a factor of $(1 + \beta)$, so that[3] $p_{i1} = p_{i0}(1 + \beta)$. Firms only offer two types of contracts: skilled and unskilled. Workers receive offers from firms with a positive probability (regardless of their employment status) and leave employment at an exogenous rate of job destruction. Workers have a reservation value b. We consider (as in the original model) that workers can pay for their training and that they make a payment that is exactly equal to the amount of benefits they will recoup in the future from being skilled (in the Appendix, we provide a detailed proof showing that the results do not change if we eliminate this assumption). This trick implies that value functions for workers only depend on the wage paid: An employer offering a higher wage is always preferable to an employer offering a lower wage because the worker is indifferent between being trained or not. Therefore, the probability that a worker separates from a firm (separation rate - s) depends negatively on the wage, given the distribution of wages in the economy, F, while the probability that a worker accepts an offer (recruitment rate - R) depends positively on the wage, for any skill level ($T = 0, 1$). Workers maximize their utility (the detailed derivation is provided in the Appendix). The firm takes these value functions in its profit maximization problem and solves:

$$\max_{w_{i0};w_{i1};T} \pi_i = (p_{i0} - w_{i0})N_{i0} - (p_{i0}(1 + \beta) - w_{i1})N_{i1} - [c - (V_{i1} - V_{i0})]T_i \tag{2}$$

In the steady-state, each firm i employs a constant number of skilled and unskilled workers,

[2] A linear technology in unskilled and skilled labor is admittedly a simplification, but one that is widely used throughout this literature (Burdett and Mortensen, 1998).

[3] Even though the training technology is homogeneous across employers in Brazil, since training is provided by SENAI, the same skills are likely to impact productivity differently depending on the type of firm. For instance, even if all workers learn how to operate a computer, training would have different effects on production depending on the intensity with which each firm uses computers.

which implies that:

$$N_{i0} = \frac{R_{i0}(w_{i0}; F_0) - T_i}{s_{i0}(w_{i0}; F_0)} \tag{3}$$

$$N_{i1} = \frac{R_{i1}(w_{i1}; F_1) + T_i}{s_{i1}(w_{i1}; F_1)} \tag{4}$$

Given 3 and 4, one can rewrite 2 as:

$$\begin{aligned} \max_{w_{i0};w_{i1};T} \pi_i =& (p_{i0} - w_{i0})\frac{R_{i0}(w_{i0}; F_0) - T_i}{s_0(w_{i0}; F_0)} + (p_{i0}(1+\beta) - w_{i1})\frac{R_{i1}(w_{i1}; F_1) + T_i}{s_{i1}(w_{i1}; F_1)} \\ & - [c - (V_{i1} - V_{i0})]T_i \end{aligned} \tag{5}$$

The first-order conditions of this problem imply[4]:

$$w_{i0} = \frac{\varepsilon^R_{i0} - \varepsilon^S_{i0}(1-\theta_{i0})}{1 + \varepsilon^R_{i0} - \varepsilon^S_{i0}(1-\theta_{i0})} p_{i0} \tag{6}$$

$$w_{i1} = \frac{\varepsilon^R_{i1} - \varepsilon^S_{i1}(1+\theta_{i1})}{1 + \varepsilon^R_{i1} - \varepsilon^S_{i1}(1+\theta_{i1})} p_{i0}(1+\beta) \tag{7}$$

$$[\frac{p_{i0}(1+\beta) - w_{i1}}{s_{i1}} + V_{i1}] - [\frac{p_{i0} - w_{i0}}{s_{i0}} + V_{i0}] - c = 0 \tag{8}$$

where w_{i0} is the wage of unskilled (not trained) workers, w_{i1} is the wage of skilled (trained) workers, ε^R_T is the elasticity of the retention rate to the wage, ε^S_T is the elasticity of the separation rate to the wage, $\theta_{i0} = \frac{T_i}{R_{i0}}$ (the share of recruited unskilled workers that the firm trains), $\theta_{i1} = \frac{T_i}{R_{i1}} = \theta_0 \frac{R_{i0}}{R_{i1}}$ (ratio of trained workers to skilled workers recruited externally) and c is the cost of providing training (which we assume to be homogeneous, for simplicity). $V_{i0}(w_{i0})$ and $V_{i1}(w_{i1})$ denote the value of being a skilled worker and the value of being an unskilled worker at the firm.

These wage-setting conditions allow us to derive a set of propositions:

Proposition 1 *Coeteris paribus, the higher the productivity level of a firm, p_{i0}, the higher the skilled and unskilled wage paid by the firm.*

This result is implied by the fact that: $\frac{w_{i0}}{p_{i0}} > 0$ and $\frac{w_{i1}}{p_{i1}} > 0$.

Proposition 2 *As long as the recruitment and separation elasticities are finite, firms always pay below marginal productivity. Furthermore, both skilled and unskilled wages increase if elasticities of recruitment and separation increase in absolute value. This implies that firms with higher market power will pay lower wages, for the same productivity levels.*

[4]The derivation and interpretation of these conditions are provided in the Appendix.

Proposition 2 is very intuitive: The more sensitive the recruitment and separation rates to changes in the wage, the higher the incentives for the firm to increase wages. This can be seen in Conditions 6 and 7 since the numerator is always smaller than the denominator, as long as the labor supply is not perfectly elastic.

From now on, as in Manning (2013), we assume that the employer chooses w_0 and w_1 such that they represent the same point in both wage distributions, F_0 and F_1. Denote this point by F. Given Proposition 1, and analogously to what is shown for the model with no training in Bontemps et al. (2000), the distribution of wages will be a continuous and increasing function of the distribution of productivities in the same support. This means more productive firms will be positioned further up in the distribution of both unskilled and skilled wages.[5] Since the recruitment and separation rates only depend on F (if all firms pay above the unemployment benefit b), we can rewrite the firm shares of employment, N_0 and N_1, as:

$$N_{i0} = \frac{R_i(F) - T_i}{s_i(F)} \tag{9}$$

$$N_{i1} = \frac{R_i(F) + T_i}{s_i(F)} \tag{10}$$

and the total employment share as:

$$N_{i0} + N_{i1} = 2\frac{R_i(F)}{s_i(F)} \tag{11}$$

From these conditions, and keeping in mind Proposition 1, which implies that higher productivity results in higher recruitment rates and lower separation rates, we can derive the following result:

Proposition 3 *Firms with relatively higher productivity levels (p_{i0}) will have higher market shares inside the relevant labor market.*

Lastly, we are interested in predicting the distribution of training across firms. Again, heterogeneity in productivity levels is not a necessary assumption for a distribution of wages and market shares in Burdett and Mortensen's seminal model. However, it allows us to pinpoint the ordering of firms in the final distribution of wages and make a prediction regarding the relationship between market shares and training rates. To do this, we turn to the last first-order condition, Equation 8, from which we can rewrite the marginal benefit of training as:

$$\frac{\beta p_{i0} - (w_{i1} - w_{i0})(F)}{s(F)} + (V_{i1} - V_{i0}) \tag{12}$$

[5] Mathematically: $p_{iT} > p_{jT} \Rightarrow w_{iT} > w_{jT}$ for $i \neq j, T = 0, 1$

Proposition 4 *The higher the productivity gains from training, which depend on initial productivity levels* (p_{i0}) *and direct productivity returns to training* (β)*, the higher the incentives to train, as long as there is wage compression, meaning* $\beta p_{i0} > (w_{i1} - w_{i0})$.

Proposition 4 suggests that we should expect to observe higher levels of training in better firms. Moreover, it justifies why the data usually does not show that future trainees receive lower wages initially, when compared to workers in firms that do not train: in fact, they receive lower wages the higher the rate of training at the firm (see Condition 6) but they receive a wage premium from working in better firms. The assumption of wage compression ($\beta p_0 > (w_1 - w_0)$) is realistic: The empirical evidence points that wage increases tend to be lower than increases in productivity arising from training (see, for instance, Barron et al. (1997)). In fact, it is enough that labor markets for skilled workers are less competitive than labor markets for unskilled workers for wage compression to arise in equilibrium. Furthermore, we are assuming that training has a proportional effect on productivity in all firms. An extension of this model could consider heterogeneity in the effectiveness of training - or skill bias incorporated into the production function - which would further increase the incentives to train in better firms.

All in all, this model predicts a positive correlation between market shares and training rates. The channel is heterogeneity in differences in productivity: More productive firms pay higher wages which allow them to recruit and retain a higher share of workers. Moreover, more productive firms also have more incentives to train because the returns to training are relatively higher as well.

3 Institutional Context and Data

3.1 Institutional Context

Serviço Nacional de Aprendizagem Industrial (SENAI) is the largest training provider in Brasil. It is a privately-run not-for-profit organization established by the Brazilian Confederation of Industry (a patronal syndicate). It is run by 27 regional departments and has more than 800 operating units. SENAI is partly funded by a mandatory contribution by firms, which amounts to 1% of the total wage bill,[6] and allows firms to have access to training courses provided regionally. Although the system is legally bound to channel two-thirds of the revenue coming from the compulsory contribution to the provision of free training courses, it also provides other courses where workers, and firms, can enroll by paying a fee. Overall, SENAI revenues amount to over 1 billion dollars every year.[7]

[6] Firms with more than 500 employees pay an additional contribution of 0.2%.

[7] In 2017, the year of the earliest budget report we got access to, the total amount of revenues was 5.615.024.386 reais ($\approx$ 1.13 billion dollars), from which 3.224.259.211 reais ($\approx$ 648 million dollars) were raised from compulsory contributions.

The Brazilian training setting provides a unique opportunity to study extensively the employer's decision towards training, particularly for two factors that set it apart from other training programs. Firstly, training programs are not tailored for a specific firm, but rather designed to meet the demand for skills in an occupation or region. Therefore, we can study what is usually regarded in the literature as general training: investment in skills that are transferable within firms that employ workers in the same occupation. This aspect is particularly relevant to establishing a link between the provision of training and market power, since all firms operating in the same market may benefit from this investment. This contrasts with the case of most countries, where job training programs are decentralized and thus very heterogeneous across firms.

Secondly, contrary to what usually happens, a large part of the cost of training becomes sunk since it is independent of the amount of training provided and is paid beforehand. This reduces and homogenizes the *per-worker* training costs that firms face. Arguably, there is a part of the cost that is still conditional on the amount of training provided even if no fees are due: For instance, the forsaken production due to lower productivity during training or administrative costs from signing up for training and dealing with paperwork. Nevertheless, we can see the upfront payments to SENAI as a significant reduction in the individual direct expense of training. This allows us to place a special focus on the most relevant market-specific cost the employer faces when making a training decision: The threat of employee turnover, which hinders the firm's ability to recoup its investment in training.

Although SENAI is responsible for about 80% of Brazil's training Blyde et al. (2019), one could argue that measuring training based on these administrative records only ignores other forms of employer-sponsored training provided inside the firm. However, the existence of these programs would likely bias our results downwards. First, larger firms are more likely to have the financial capacity to promote their own training programs. Assuming that own training courses crowd out SENAI training, not observing these programs would result in a negative correlation between market power and training provision. Furthermore, this would imply that we are including in the control group some groups of workers that are actually trained as well, which would bias downwards the estimated effect on wages.

3.2 Data

This paper uses two main administrative datasets from Brazil: the *Relação Anual de Informações Sociais* (RAIS) longitudinal census containing yearly information at the employer-employee-job title level and the *Serviço Nacional de Aprendizagem Industrial* (SENAI) administrative records including individual information on the workers trained each year.

SENAI

The SENAI dataset ranges from 2009 to 2012 and includes information on the individuals trained each year, the municipality of the training facility, the modality of the course taken, the course duration, and the enrollment and completion dates.

SENAI provides five types of job training: (1) technical upgrading; (2) apprenticeship; (3) initiation; (4) habilitation; and (5) qualification courses. We chose to analyze technical upgrading courses[8] as their goal differs substantially from the goal of other types of courses. The purpose of a technical upgrading course is to provide a relatively short (fewer than 160 hours, but on average 40 hours) course that aims at updating or complementing the professional skills of a specific occupation or field. Therefore, the curriculum is very job-specific, and is oriented towards updating workers' skills to meet technological advancements or to deal with new processes in their occupation. Some examples of technical upgrading courses are *"Technical Upgrading on Reading and Interpreting Mechanical Design"* or *"Technical Upgrading in preparing Italian risottos and pasta"*. This means that the goal of these courses is not to introduce a worker to a new job (like apprenticeship, initiation, or habilitation courses) or to change their career track (like qualification courses) but rather to expand already acquired professional skills.

SENAI courses are open to the general public, implying that a worker does not need to be employed to complete a training course. However, given the nature of technical upgrading courses, more than 95% of the students completing these courses do so through their employer.

RAIS

RAIS is collected by the Ministry of Labor in a compulsory survey to all formal firms and their registered workers. We use these records for the 2009-2013 period. RAIS provides detailed information at the worker level (such as age, gender, schooling, and race) and at the establishment level (municipality and industry) as well as job characteristics (occupation, wage, hours worked, tenure, hiring and termination dates, and type of contract). Several cross-year checks were performed to guarantee the consistency of the data, particularly regarding gender, education, and age. We keep in our final sample full-time workers aged 16-65 with non-missing occupation and wage, preserving more than 99% of the initial dataset. RAIS offers information for each match worker-establishment. This allows us to perfectly trace the number of courses a worker undertakes while employed in each establishment by crossing training dates with hiring and firing dates.

[8]From 2010 onwards, SENAI reclassified basic qualification courses with a duration of fewer than 160 hours as technical upgrading courses, regardless of the content of the course. This creates an inconsistency between 2009 and 2010-12. To ensure consistency, and following the most recent definition used, we reclassified all qualification courses of 2009 with less than 160 hours as technical upgrading courses.

Sample Construction

The focus of this thesis is the firm decision to promote training. Therefore, our unit of interest is not the employee, but rather the employer, the firm. This thesis relies on an important distinction between three types of *"firm-level"* units: (i) the firm; (ii) the establishment; and (iii) the establishment-occupation. The firm is the most aggregated of these three units and is composed of one, or more, establishments.[9] Each establishment employs several groups of workers with different occupations, which we call *establishment-occupation groups*. Given the job-specificity of the training courses examined, and since we focus on labor market dynamics, this latter unit is our main unit of analysis. To illustrate what an establishment-occupation group is, consider, for instance, an establishment called *Supermarket 1*. Our main unit of analysis is not the establishment itself, but rather each group of workers with a different occupation working in this supermarket. Specifically, we can have *Technical Operators of Supermarket 1*, *Merchants and Sellers of Supermarket 1*, and *Maintenance and Reparation Workers of Supermarket 1*. To construct this dataset, we leverage the worker dataset previously built after crossing RAIS and SENAI to compute the measures of interest (such as the average wage, number of trainees, etc...) for each establishment-occupation group of workers. All measures of stocks (such as the number of employees) refer to the stocks as of 31^{st} of December of the year. We do, however, compute measures of flows (such as the number of displaced workers) throughout the whole year. To classify occupations, we use the Brazilian occupation classification (CBO - *Classificação Brazileira de Ocupação*) at the two-digit level. More detail on this classification can be found in Table 13 in the Appendix.

We work at the two-digit level of sectors following the Brazilian classification (we use CNAE 2.0 - *Classificação Nacional de Atividades Económicas*), for which more detail is provided in Table 11 in the Appendix. It is worth noticing that SENAI training courses primarily focus on industry, since there is another organization (**SENAC** - *Serviço Nacional de Aprendizagem Comercial*) that handles training in the commercial services sector. With this in mind, as an alternative to dropping all firms operating in the services sector (which could be misleading since there are firms operating in more than one sector and there are some SENAI courses directed at specific services, such as transportation), we decided to only keep in our final sample establishments that belong to sectors where at least one firm provides training during the four-year period.[10] Similarly, we only keep in our final sample establishment-occupation groups operating in municipalities and occupations where there is at least one training course during the same period.

[9]The large majority of the firms in our dataset operate one establishment only.

[10]This means that the sectors of Agriculture, Real Estate, Public Administration, Social Services, Domestic Services, and International Organizations are excluded.

3.3 Descriptive Statistics

Although our main analysis is at the establishment-occupation level, for which we present detailed statistics later in this section, we start by introducing some descriptive statistics at the establishment level. Table 1 reports the categories used by SEBRAE (*Serviço Brasileiro de Apoio às Micro e Pequenas Empresas*) to classify establishments according to the number of employees. We use the same classification in this thesis, and present all estimates also for the sub-sample excluding microestablishments[11]. In Table 2 we report the mean and standard deviation of some key variables in establishments where at least one worker is trained in the period 2009-2012 and in establishments where no worker is trained in the same period. Establishments that train are significantly larger, employ younger and more educated workers, and seem to have a higher average worker tenure.

Table 1: Classification of establishments according to the number of employees.

Classification	**Industry**	**Services**	**% in data**
Microestablishment	0-10	0-20	82%
Small establishment	10-50	20-100	15%
Medium Establishment	50-100	100-500	2%
Large Establishment	100+	500+	1%
All	12%	88%	100%

Note: This table presents the classification of each establishment according to the number of employees following the definition of size used by *SEBRAE*.

Table 3 presents some descriptive statistics at the establishment-occupation level for the final sample used in the remaining of this thesis, which includes establishment-occupation groups that besides meeting our sampling criteria, are not singleton observations.[12] In the Appendix, the same statistics for the sample excluding microestablishments are provided in Table 9. A small note should be made with regards to the proportion of units treated: Comparing the number of observations in Tables 2 and 3, the proportion of trained units falls from almost 7% to approximately 5.3%. This is expected, since it is enough that a worker in one occupation is trained for the establishment to be considered "treated". However, once we split each establishment in different occupations, some of these units may shift to the non-treated group. In the Appendix, Figure 6 displays the number of trainees, as well as the training rate, by sector. Table 10 presents the number of unique firms, establishments, and establishment-occupation groups included in our analysis.

[11]We present all estimates also for the sample excluding microestablishments since a large part of these establishments are family-owned businesses that do not necessarily behave similarly to large competitive firms. We show, therefore, that our results are robust to considering relatively larger firms only.

[12]This implies excluding firms with only one observation at the establishment-occupation level during the whole period of analysis.

Table 2: Descriptive Statistics for establishments

	Trains at least once		Never Trains		
	Mean	S.D.	Mean	S.D.	Difference
Number of employees	116.899	1337.288	7.542	34.844	-109.357***
Average Real Wage	93,435.134	79,494.451	62,809.535	53,891.072	-30,625.598***
Average employee age	33.439	5.798	33.833	8.678	0.394***
Average employee tenure	3.081	2.880	2.805	3.320	-0.276***
Share of women employees	0.323	0.284	0.480	0.407	0.157***
Share of employees with 9+ years of education	0.872	0.199	0.870	0.273	-0.002***
Share of employees with 12+ years of education	0.625	0.309	0.617	0.404	-0.009***
Share of employees with a university degree	0.087	0.167	0.066	0.191	-0.021***
Number of trained employees	1.990	19.590	0.000	0.000	-1.990***
Observations	755,943		10,980,679		

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: This table presents the average and standard deviation of some variables of interest for establishments where at least one worker is trained in the period 2009-2012 and for establishments where no worker is trained in the same period. Real wages are obtained from annual nominal wages which we deflate using a consumer price index for Brazil.

Table 3: Descriptive Statistics for establishment-occupation groups

	Trains at least once		Never Trains		
	Mean	S.D.	Mean	S.D.	Difference
Number of employees	48.834	695.730	4.116	22.154	-44.719***
Average Real Wage	115,887.967	114,610.874	83,865.033	112,100.844	-32,022.934***
Average employee age	33.665	7.211	34.498	9.772	0.833***
Average employee tenure	3.631	3.813	3.124	3.939	-0.506***
Share of women employees	0.252	0.328	0.457	0.453	0.206***
Share of employees with 9+ years of education	0.883	0.220	0.885	0.281	0.002***
Share of employees with 12+ years of education	0.661	0.351	0.655	0.430	-0.005***
Share of employees with a university degree	0.107	0.249	0.111	0.292	0.004***
Number of trained employees	1.157	8.456	0.000	0.000	-1.157***
Observations	1,265,776		23,797,234		

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: This table presents the average and standard deviation of some variables of interest for establishment-occupation groups where at least one worker is trained in the period 2009-2012 and for establishment-occupation groups where no worker is trained in the same period. Real wages are obtained from annual nominal wages which we deflate using a consumer price index for Brazil.

4 Labor Market Definition

Given that labor markets have not been a focus of antitrust authorities, a consensus has not been reached in the literature as to how to define the relevant market. Guiding ourselves by the criteria set in the Horizontal Merger Guidelines, which state that the definition of a (product) market should encompass both *"line of commerce"* and *"section of the country"*, we will define labor markets using two-digit-level occupation and a geographical zone delimited by the municipality where the firm is established, in line with what is done in the recent literature that studies the effect of labor market concentration on wages (Azar et al., 2022). We decided to keep the Brazilian classification of occupation rather than translating it to the *International Standard Classification of Occupations* (ISCO) for several reasons. First, it has been noted in the literature (Muendler et al., 2004) that the Brazilian occupation classification is more profession-based while the international classification is more skill-oriented. Since we can finely create a proxy for skill level using education, we consider that the Brazilian classification may help provide a more accurate definition of labor markets[13]. Secondly, as we are evaluating a training program in Brazil and SENAI vacancies are defined and designed for the occupational framework of Brazil, we consider that the Brazilian occupation classification is more relevant. Table 13, in the Appendix, provides more detail on the classification of occupation used.

Naturally, this approach has several limitations presented here. First. it relies on a rather wide definition of occupation, and disregards the role of sector or the characterization of a job. A possible next step would be to consider a more restrictive definition of labor market where we cross sector and occupation. A second assumption we are making is that most do not seek jobs outside the municipality where they work. In fact, what matters for the definition of the relevant labor market is the mobility patterns of the population. In the literature, which largely focuses on US markets, it is common to use commuting zones as the geographical delimitation of labor markets (Azar et al., 2022; Rinz, 2022), but the same zones are not defined for Brazil. Reassuringly, data from the Population Census of the same period (2010) suggests that the vast majority of the population (more than 90%) does not commute outside its municipality of residence.

Across the 5,570 municipalities, 46 different labor markets are considered each year, yielding a total of 936,438 municipality-occupation-year observations with a non-missing number of workers. In our mapping of labor market concentration across Brazil, we keep labor markets where there is at least one employee. However, for the main part of our analysis at the establishment-occupation level, where we use labor market concentration and labor market power to predict the decision to train, we exclude markets with a labor force below 10 workers.

In the following sections, two important measures of labor market power and labor market concentration are used. As a measure of market power, we compute the share of the labor market mj (a combination of municipality m and occupation j) that the firm i[14] employs in

[13]Nevertheless, after translating CBO to ISCO, we found the differences at the two-digit level to be minimal.

[14]As in Lipsius (2018), we treat all establishments belonging to a single firm operating in a common market as one single establishment. This assumption does not affect any of our results since only a very small number of firms operate two or more establishments in the same labor market.

year t:

$$s_{ijmt} = \frac{N_{ijmt}}{\sum_i N_{ijmt}} \times 100 \tag{13}$$

where N_{ijmt} denotes the number of workers from firm i, in occupation j and municipality m, in year t, and $\sum_{i=1}^{I} N_{ijmt}$ is the total labor force in the market defined by municipality m and occupation j, for year t, where I firms operate. As a baseline measure of labor market concentration, we use the *Herfindahl–Hirschman Index* (HHI), a measure widely adopted by antitrust authorities to measure the competitiveness of markets, that ranges from 0 (minimum concentration) to 10.000 (monopolistic market):

$$HHI_{jmt} = \sum_i s_{ijmt}^2 \tag{14}$$

While its simplicity makes it a very attractive measure to summarize market concentration, it has no economic meaning beyond that attributed by the thresholds commonly used by antitrust authorities (Azar et al., 2022), which classify markets with an HHI between 1500 and 2500 as ”moderately concentrated” and above 2500 as ”highly concentrated”. Figure 1 reports the average HHI in each Brazilian municipality, where occupations are weighted by the relative size of their labor force.

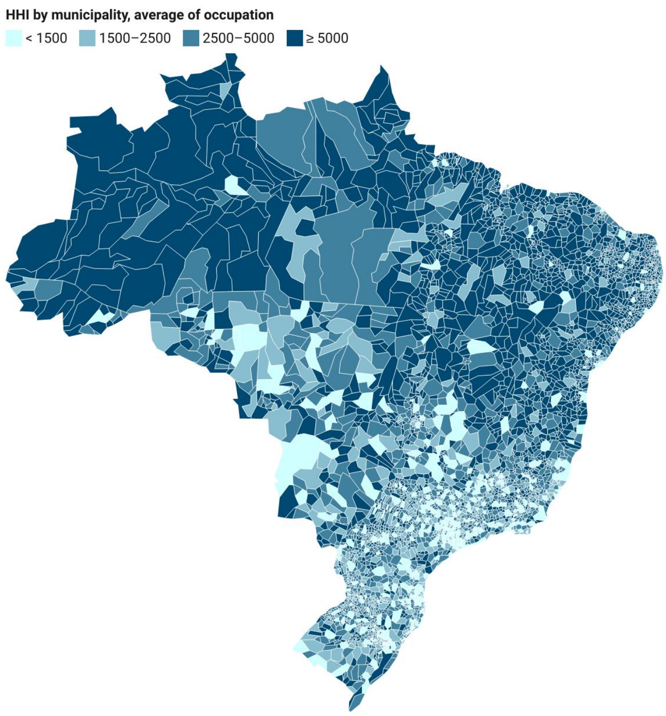


Figure 1: Labor market concentration in Brazil by municipality

Note: The map presents the HHI calculated for each Brazilian municipality, as an average of the HHI of each occupation, weighted by the labor force employed in each occupation. We use the same cut-off points for categories as Azar et al. (2022).

5 Drivers of Training

Before turning to the main estimates of the effect of training on key outcome variables, we conduct an analysis of the relationship between labor market concentration, labor market power, and training.

5.1 Empirical Strategy

In this section, we study the drivers behind the firm decision to train. We measure training provision with a binary variable ($Train_{ijmt}$) that takes the value of one if at least one worker in the establishment-occupation group considered undergoes training during the year, and zero otherwise:

$$Train_{ijmt} = \begin{cases} 0, & \text{if no worker is trained from the group } ijm \text{ in year } t \\ 1, & \text{if at least one worker is trained from the group } ijm \text{ in year } t \end{cases} \tag{15}$$

where i denotes the firm, j is the occupation group, m denotes the municipality of the establishment and t corresponds to the year of analysis. This binary measure of training splits establishment-occupation groups into those who promote some type of training and those who do not, ultimately capturing engagement in training.

We estimate, using ordinary least squares estimation, a linear probability model of the form:

$$Train_{ijmt} = \mathbf{X}'_{1ijmt}\alpha + \mathbf{X}'_{2ijmt}\beta + \gamma_{mj} + \theta_{st} + \varepsilon_{ijmt} \tag{16}$$

where $\mathbf{X}_{1ijmt}$ is a variable, or a set of variables, that characterize the market power of firm i in labor market jm in year t. For instance, in our main specification, this variable is the natural logarithm of the market share of the establishment-occupation group ijm in market jm. $\mathbf{X}_{2ijmt}$ is a set of establishment-occupation level controls that includes the number of employees and its square, average employee age and tenure, share of workers with 9+ and 12+ years of education, and share of women. We include these variables to account for differences between establishment-occupation groups that impact training take-up and could bias our final results if correlated with measures of market concentration. Here, special focus should be placed on the inclusion of firm size and its square. As established in Section 3, larger units have a higher probability of training at least one worker. However, market power does not depend on absolute size only, but also on the number of firms operating in the same market and the relative size of each group. Properly accounting for the effect of absolute size allows us to estimate the impact of market share "cleaned" of the pure effect of size. Lastly, γ_{mj} and θ_{st} denote municipality $\times$ occupation and sector $\times$ year fixed effects, respectively.

In some specifications, ϕ_i, firm fixed effects, are included. Including these fixed effects controls for unobserved factors common to all establishments of the same firm. For instance, they may capture firm-specific differences in productivity, or a common wage-setting policy. Typically, the decision of the clustering level is *"motivated by the concern that unobserved components of outcomes for units within clusters are correlated"* (Abadie et al., 2017). In this case, we consider that unobserved shocks that may affect training and create correlation in our standard errors are likely to happen at the labor market level. Therefore, we cluster standard errors at the municipality $\times$ occupation level.

Here, and in the remaining of this thesis, it will become clear that, although we refer to market concentration and control for it using the *Herfindahl–Hirschman Index* in some of our specifications, we focus our attention mainly on market power measured by market share. The reason for this is the following: Our empirical strategy, and particularly, the validity of the instrumental variable we use, relies heavily on the inclusion of municipality-occupation (labor-market level) fixed effects. Therefore, the identification of the effects of market concentration would be done solely through inter-annual variations in concentration inside each market, which, given the relatively small time-span of our analysis, are relatively limited.

5.2 Results

Table 4 presents the estimation results from Equation 16. We capture market power using four different measures to show that the results are robust to different definitions. First, we use the log of the establishment-occupation share in the relevant labor market (s_{ijmt} to capture the establishment's market power, to which we add the log of the market HHI (HHI_{jmt} to control for differences in market concentration. Then, we compute two alternative measures of market power to make interpretation easier: A binary variable, *Largest Firm*, which takes the value of one if the establishment employs the highest share of workers in its market, for the relevant occupation (meaning, is the firm with highest market power); and a categorical variable with the quartiles of market share (computed after ordering the establishment-occupations groups inside the relevant market). We also add firm fixed-effects to our preferred specification. Aligned with the theoretical predictions, we find that market share and the market HHI are important predictors for the establishment decision to train a specific occupation.

Columns (1) and (6) show that, if we compare two establishments with the same observable characteristics, accounting for municipality and occupation-specific effects, a 1% higher market share means, on average, a 3 percentage point (p.p.) higher probability that the firm trains. This estimate is robust to controlling for market concentration as columns (2) and (7) show. Furthermore, these columns suggest that two establishments with the same observable

Table 4: Establishment-Occupation level determinants of Training, OLS estimates

	Full sample					**Excluding Microestablishments**				
Dependent Variable: $Train_t$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$Ln(share_t)$	0.026***	0.026***			0.024***	0.035***	0.036***			0.032***
	(0.001)	(0.001)			(0.001)	(0.001)	(0.001)			(0.001)
$Ln(HHI_t)$		0.004***					0.006***			
		(0.000)					(0.000)			
Largest Firm			0.096***					0.102***		
			(0.002)					(0.002)		
2^{nd} Quartile				0.001***					-0.003***	
				(0.000)					(0.001)	
3^{rd} Quartile				0.009***					0.010***	
				(0.000)					(0.000)	
4^{th} Quartile				0.054***					0.054***	
				(0.001)					(0.001)	
Observations	25,063,010	25,063,010	25,063,010	25,063,010	25,063,010	10,580,853	10,580,853	10,580,853	10,580,853	10,580,853
R-squared	0.085	0.085	0.056	0.073	0.205	0.114	0.114	0.077	0.198	0.212
Year X 4-digit sector	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	No	No	Yes	No	No	No	No	Yes

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: This table displays the outcomes of estimating a regression whose dependent variable is equal to one if the firm trains at least one worker and zero otherwise with fixed effects estimation. Columns (1)-(5) show the effect of market power and market concentration on the probability of training using all establishment-occupation groups in labor markets with at least 10 employees. In columns (6)-(10), groups that belong to microestablishments are excluded from the sample. The variable "Largest Firm" is an indicator variable that is equal to one if the establishment-occupation group is the one with the highest share in its market and zero otherwise. In columns (4) and (9), quartiles are calculated by ordering raw market shares for each labor market. Establishment-occupation controls include the number of employees and its square, share of women, average worker age and tenure, and share of employees with 9+ and 12+ years of education. Standard errors clustered at the municipality $\times$ occupation level in parentheses.

characteristics including the same share and size will differ in levels of training as long as they belong to markets with different concentration levels. In particular, an establishment located in a more concentrated labor market will train more than an establishment with similar characteristics located in a more competitive market. One might worry that including different measures for market concentration in the same regression (such as market share and the HHI) may create a multicollinearity issue and threaten our ability to properly identify each coefficient. We argue, however, that this is not the case: Firstly, the standard errors are relatively small in all specifications, even when more than one variable is included. Secondly, the coefficients do not seem to be sensitive to the inclusion of other covariates - for instance, the estimated effect of market share does not differ significantly from column (1) to (2), where a measure for the HHI is included, and are exactly equal in columns (6) and (7).

Columns (3)-(4) and (8)-(9) confirm that the relative size of an establishment inside a labor market is a significant determinant of the firm incentives to train. Specifically, being the largest establishment in a labor market translates into a 10 percentage point higher probability of training, when compared to the average for the other establishments. Moreover, on average, the probability of training is 5 p.p. higher among firms in the fourth quartile of labor market share, when compared to firms in the first quartile. Lastly, the inclusion of firm fixed effects changes only slightly the estimated impact of market power on the probability of training. In fact, even when we compare establishment-occupation groups of the same firm, we observe a significantly higher probability of training among units with higher market power.

To provide a visual sense of the relationship between labor market power and training, we estimate the average probability that an establishment-occupation group is trained as a function of its rank inside the labor market where it belongs, when units are ordered by (residualized) market shares (following the methodology used in Groes et al. (2015)). We generate residual market shares by estimating a standard reduced-form regression of market share on a set of establishment-occupation controls ($\mathbf{X}_{ijmt}$) that include the number of employees and its square, share of women, average worker age and tenure, and share of employees with 9+ and 12+ years of education:

$$share_{ijmt} = \alpha + \mathbf{X}'_{ijmt}\beta + \varepsilon_{ijmt} \tag{17}$$

The residuals we use to create the ranks of market power, $\hat{\varepsilon_{ijmt}}$, represent the part of the market share of the establishment that is not explained by observable factors. For instance, the market power of an establishment is partly determined by its size. However, two establishments of the same size in different markets may differ considerably in their market power depending on the size, and structure, of the market in which they operate. Figure 2 displays the predicted probability that firms in a given decile of their labor market train (averaging on the other covariates), computing these average predicted values separately for five different levels of market concentration. Figure 2, therefore, presents three dimensions: the probability of training (y axis), relative market power (x axis), and market competitiveness (different curves).

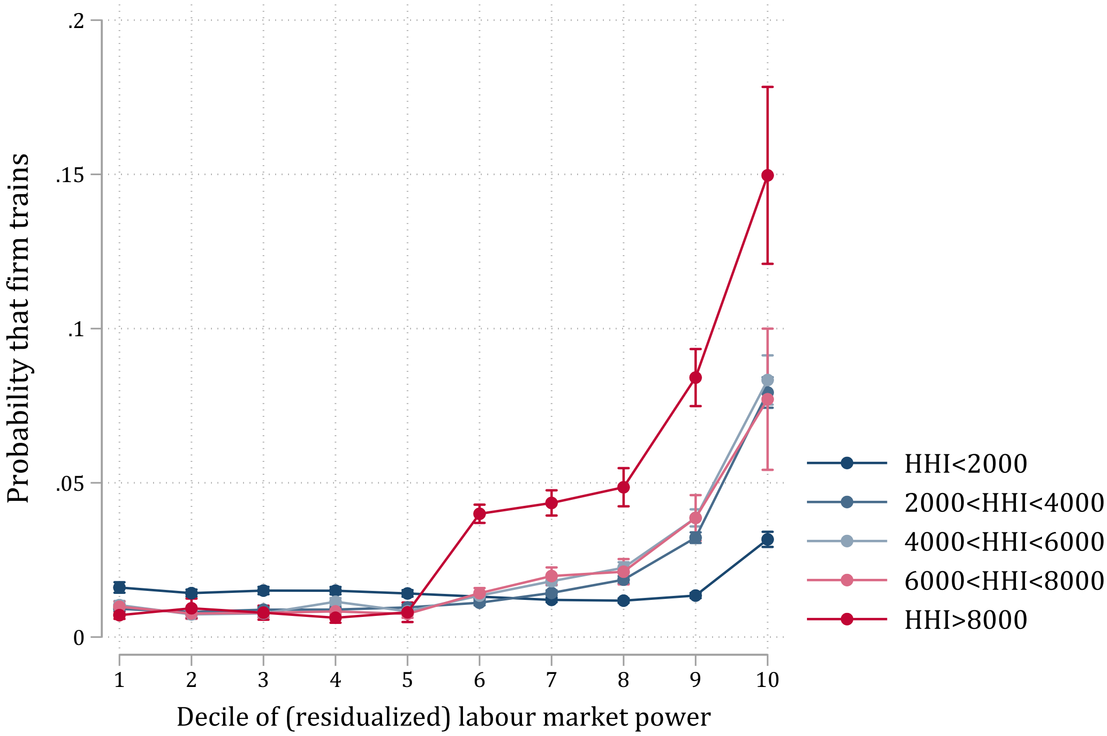


Figure 2: The relationship between market power and training, by market concentration

Note: The graph displays the margins of a regression of an indicator variable equal to one if the establishment-occupation group has at least one worker that is trained and zero otherwise on the residualized decile of labor market power in the relevant labor market. This residual is obtained from a regression on the number of employees and its square, share of women, average worker age and tenure, and share of employees with 9+ and 12+ years of education. Standard errors are clustered at the occupation $\times$ municipality level.

This plot provides two main insights into the relationship between training and market power: On the one hand, aligned with the results presented in Table 4, we find that the probability of training is highest among the largest firms in each market, regardless of the overall level of market concentration: Even in competitive labor markets, bigger firms tend to train more than smaller ones. On the other hand, the graph reveals that the effect of *"being big"* (a higher decile) is larger in more concentrated markets, suggesting that, together with the relative rank inside each market, the *absolute* value of the market share of a firm matters. The nonexistent or even negative relationship between HHI and training for lower deciles (particularly, deciles 1-5) also hints that the positive correlation detected between market concentration and prevalence of training is largely driven by the fact that firms in more concentrated markets tend to have higher market shares.

If we go back to our theoretical framework, we understand that the positive relationship between market power and training is not surprising. In our model, firms gain market share by paying higher wages, which reduces the likelihood that the worker separates from the firm (an important variable in the decision to train). If this is true in reality, then we should expect to see a negative correlation between market power and separation rates. To test this, we next turn to the relationship between market power and the separation rate. We measure the latter using a definition close to the one used in Davis and Haltiwanger (1990), but including only the number of workers that separated from the firm during the year[15] (instead of both separations and ascensions) divided by the size of the establishment-occupation group, measured by the average number of employees in the relevant establishment-group:

$$Separation\ Rate = \frac{Separated_{ijmt}}{(N_{ijmt-1} + N_{ijmt})/2} \tag{18}$$

We estimate a regression of the job separation rate on market share at the establishment-occupation level and a set of establishment-occupation controls. Figure 3 displays the predicted values of the separation rate for different levels of market share, averaging on the remaining controls. As expected, we find a strong negative correlation between employee separations and market power in the data: A 10 percentual point increase in market share corresponds, on average, to a 2 p.p. decrease in separations. The same is true at the market level: As Figure 4 displays, the more competitive the market, the higher the number of separations.

[15] For a worker to separate from firm i, it could either be because the worker is employed at a new firm, unemployed or employed in the informal labor market.

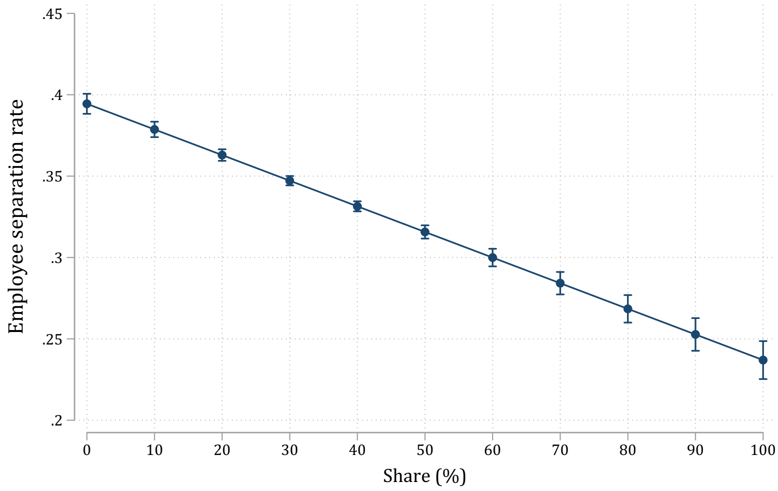


Figure 3: Separation Rate and Market Share

Note: The graph displays the predicted separation rate for different values of market share, as well as confidence intervals at a 95% confidence level, resulting from a regression where we also control for the number of employees and its square, share of women, average worker age and tenure, and share of employees with 9+ and 12+ years of education. Standard errors are clustered at the occupation × municipality level.

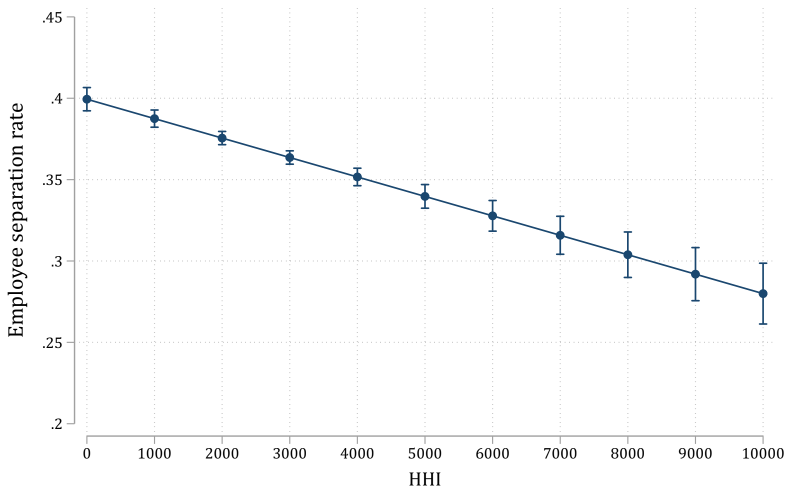


Figure 4: Employee Separation and Market Concentration

Note: The graph displays the predicted separation rate for different values of market concentration (measured by labor market HHI), as well as confidence intervals at a 95% confidence level, resulting from a regression where we also control for number of employees and its square, share of woman, average worker age and tenure and share of employees with 9+ and 12+ years of education. Standard errors are clustered at the occupation × municipality level.

6 Effects of Training

After establishing a link between market power and incentives for training, we turn to the question *"Does training affect the employer distribution of wages?"*. Assuming that training positively affects productivity, the theory predicts that firms that train have a wage structure different from firms that do not train. In particular, if, as in our model, productivity dispersion is the underlying mechanism for different training rates, then we should expect higher average wages for firms that train more.

In this section, we start by pinpointing the effects of training for workers, namely on average wage levels and wage dispersion. Following this, and motivated by the recent literature reporting a negative effect of concentration on wages (Azar et al. (2022), Benmelech et al. (2022)), we investigate whether training will reflect differently on wages depending on the structure of the labor market.

6.1 Instrumental Variable

The main challenge in identifying a causal relationship between training and returns for the workers is to control for selection into training. In fact, the establishment's decision to send a worker to training is likely to be correlated with unobservable characteristics that, in turn, drive differences in the outcomes. The OLS results may, therefore, present an upward or downward bias. For instance, a more innovative firm may seek more training for its employees and simultaneously pay higher wages. In this case, without controlling for this bias, one would overestimate the returns to training.

To address this issue, and following an approach similar to the one in Blyde et al. (2019), we instrument training with a measure of variation in the availability of training in each labor market over time. The idea behind the choice of this instrument is that it becomes more likely that an establishment sends a worker to training if it is located closer to a larger supply of courses relevant to the occupation. We measure the availability of courses by dividing the number of slots for technical upgrading courses in a given occupation within 50 miles by the number of workers in the same radius. Since we do not have access to the exact location of the establishment or of the training center, we consider a course or a worker to be within 50 miles if the center of the municipality where the course is provided is at a distance lower or equal to 50 miles from the center of the municipality where the establishment is located. Figure 7 in the Appendix provides a visual example of the calculation of the instrument. We winsorize our instrumental variable to take values between the 1^{st} and the 99^{th} percentile[16]. Reassuringly,

[16] We winsorize our instrument to reduce the influence of outliers in the predicted values of the first regression. Particularly, this procedure reduces the weight of establishments that train workers despite being located in areas with no training (possibly requiring that someone from the training center is sent to the firm or even recurring to

our first-stage results are robust even when using the raw data. Our instrument varies at the establishment-occupation level since we exclude the slots used by the establishment and the workers employed in it. The first stage equation is given by:

$$Train_{ijmt} = \alpha Z_{ijmt} + \mathbf{X}'_{ijmt}\beta + \gamma_{mj} + \theta_{st} + \varepsilon_{ijmt} \tag{19}$$

where Z_{ijmt}, our instrument, is given by:

$$Z_{ijmt} = \frac{\sum_{i \neq \tilde{i}} C_{\tilde{i}jmt}}{\sum_{i \neq \tilde{i}} N_{\tilde{i}jmt}} \tag{20}$$

where $\sum_{i \neq \tilde{i}} C_{\tilde{i}jmt}$ denotes the sum of all courses provided in occupation j within 50 miles of municipality m in year t, excluding those used by firm i. Similarly, $\sum_{i \neq \tilde{i}} N_{\tilde{i}jmt}$ denotes the sum of all workers in labor market jm in year t that are not employed by firm i. X_{ijmt} is a set of establishment-occupation level controls that includes the number of employees and its square, average employee age and tenure, share of workers with 9+ and 12+ years of education, and share of women. γ_{mj} and θ_{st} denote municipality $\times$ occupation and sector $\times$ year fixed effects, respectively. Figure 8 in the Appendix provides a visual depiction of the annual average number of courses provided within 50 miles of each municipality, $\sum \bar{C}_{ijm}$.

The main assumption behind our instrument is that, once we control for the observable characteristics and include fixed effects, a change in the availability of training only affects the outcome variable through the decision of training workers (the so-called *exclusion restriction*, in the literature). Here, the inclusion of the fixed effects plays a determinant role in answering the two main sources of endogeneity of the instrument that arise. Firstly, we eliminate the possibility of a spurious relationship between training and our dependent variable emerging from sector-specific unobservable time variations. For instance, an increase in shoe exports could drive both an increase in demand for training in the sector and an increase in wages. Secondly, we answer the argument that firms select their location based on the location of the training center, by accounting for time-invariable differences between labour markets. In fact, after including the controls and fixed effects, what the instrument captures is the yearly variation in the availability of courses, excluding the ones used by the firm. The validity of this instrument can be questioned if there are factors varying over time at the municipality-occupation level that affect both the supply of training courses and the outcomes of interest, such as unbalanced changes in the regional demand for a specific product, for instance. However, we argue that this is not a concern since, although yearly monitored and updated, the main projections done by SENAI for the vacancies offered are usually done for a 5-year window [17].

remote learning), which results in a negative value for the instrument.

[17] The period examined is included in the 2009-2014 medium-term projection.

Identification is only possible if there is enough variability of the instrument across regions and over time. Figure 10, in the Appendix, displays the variation in vacancies per worker within 50 miles for the occupation of *"cross-functional workers"*. We chose this occupation because it is the one in which more vacancies are filled every year. As the Figure depicts, the availability of courses varies quite randomly from year to year and between regions.

Table 5 displays the estimates of the first-stage regression for our preferred IV specification: the number of technical upgrading courses at a distance of 50 miles or less per worker, excluding courses and workers from the relevant establishment.[18] As expected, there is a significant positive relationship between the availability of courses and training take-up. This relationship is stable regardless of the sample we use and is robust to the inclusion of firm fixed effects. The F-statistic is higher than 150 in all specifications, well above the commonly accepted threshold of 10 (Stock et al. (2002)), thus indicating that the instrument also meets the *relevance criteria*.

In the following subsection, we estimate several variations of our baseline regression (21),

$$y_{ijmt} = \alpha Train_{ijmt-1} + \mathbf{X}'_{ijmt}\beta + \gamma_{mj} + \theta_{st} + \varepsilon_{ijmt} \tag{21}$$

where $Train_{ijmt-1}$ is an indicator variable measuring engagement in training in the previous year (as defined in 15), $\mathbf{X}_{ijmt}$ is a set of establishment-occupation level controls that includes the number of employees and its square, average employee age and tenure, share of workers with 9+ and 12+ years of education, and share of women. As before, γ_{mj} and θ_{st} denote municipality $\times$ occupation and sector $\times$ year fixed effects, respectively. y_{ijmt}, our outcome variable, is always defined at the establishment-occupation level. Standard errors are clustered at the municipality $\times$ occupation level. We present eight variations of this regression that include OLS and IV estimates on the full sample and on the sample excluding microestablishments. We also re-estimate the same regressions including firm fixed effects (ϕ_i).

6.2 Results

We now turn to our main results regarding (i) the effects of training on wages and their dispersion and (ii) the relationship between these returns and market power.

Effect on Average Wages

Do average wages of a group of workers in the same occupation increase when training is provided to at least some of the workers? Although positive wage returns are documented in

[18] All tables in this section report a smaller sample size. This is due to the inclusion of the lag of training. We again exclude singletons so that estimates with and without firm fixed effects rely on the same sample. The first-stage results correspond also to the same sample as used in the remaining tables.

Table 5: First-stage regression

	Full sample		Excluding Microest.	
Dependent variable: $Train_t$	(1)	(2)	(3)	(4)
Courses 50 miles / number of workers	0.176***	0.172***	0.223***	0.215***
(excluding own)	(0.013)	(0.013)	(0.018)	(0.017)
Observations	19,800,658	19,800,658	9,140,274	9,140,274
Establishment-Occupation Controls	Yes	Yes	Yes	Yes
Year $\times$ 2-digit sector	Yes	Yes	Yes	Yes
Municipality X Occupation	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Note: This table displays the first-stage estimates for our preferred IV specification. Columns (1)-(2) comprise all establishment-occupation groups included in our final sample. In columns (3)-(4), groups that belong to microestablishments are excluded. Establishment-occupation controls include the number of employees and its square, share of women, average worker age and tenure, and share of employees with 9+ and 12+ years of education. Standard errors clustered at the municipality $\times$ occupation level in parentheses.

the literature at the worker level, little is known about the aggregated-level measures. Going back to Section 2, when a larger proportion of workers undergoes training, the theory predicts that the wages of the unskilled workers will go down[19] and the wages of the skilled workers increase. Therefore, the effect on the average depends on the relative size of the two groups, and also on the relative magnitude of these two movements.

Table 6 reports the results of estimating our baseline model 21 with the natural logarithm of the average wage for the establishment-occupation group concerned as the outcome variable. Following what is common practice in the literature, we use a log-linear model for wages as it allows us to estimate the effects as percentage, rather than absolute, changes in the average wage. Besides making interpretation more convenient, evaluating percentage changes follows closely our theoretical framework, where we assume that the increases in productivity from training are proportional to initial productivity levels.

Focusing on the regressions excluding firm fixed effects (1)-(2) and (5)-(6), we observe that wages are significantly higher in establishment-occupation groups where at least one worker is trained. The OLS estimates point to a difference in average wages of about 17% between groups that are trained and their similar counterparts, or 11% when microestablishments are excluded. Since we look at real annual wages, this increase may come from two sources: the intensive margin (increase in hourly wages) and/or the extensive margin (increase in hours

[19] As the fraction of unskilled workers exiting to become skilled increases, the firm becomes less concerned with separations from the unskilled.

Table 6: Effect of Training on Average Wages, OLS and IV estimates

	Full sample				**Excluding Microestablishments**			
Dependent Variable: Ln(Average Wage)	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)	OLS (7)	IV (8)
$Train_{t-1}$	0.170*** (0.003)	0.219*** (0.081)	-0.030*** (0.003)	0.184** (0.075)	0.105*** (0.003)	0.162*** (0.061)	-0.024*** (0.002)	0.131** (0.052)
Observations	19,800,658	19,800,658	19,800,658	19,800,658	9,140,274	9,140,274	9,140,274	9,140,274
R-squared	0.617	0.186	0.825	0.146	0.667	0.239	0.814	0.185
Establishment-Occupation Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year $\times$ 2-digit sector	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality X Occupation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	Yes	Yes	No	No	Yes	Yes
First Stage F-statistic		187.4		177.6		155.5		153.6

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Note: This table displays the relationship between training and average wages at the establishment-occupation level. Columns (1)-(4) comprise all establishment-occupation groups included in our final sample. In columns (5)-(8), establishment-occupation groups that belong to microestablishments are excluded. Establishment-occupation controls include the number of employees and its square, share of women, average worker age and tenure, and share of employees with 9+ and 12+ years of education. The F-statistic corresponds to the Kleibergen-Paap Wald weak identification test. Standard errors clustered at the municipality $\times$ occupation level in parentheses.

worked).

Interestingly enough, the IV estimates are slightly above the OLS estimates (though not statistically different), which suggests that OLS consistently underestimates the effects of training. In the literature that estimates returns to schooling, and also for training, this finding is quite common (Blyde et al., 2019; Card, 1999, 2001). Several explanations have been discussed in the literature: Firstly, there may be an omitted variable that is positively correlated with training take-up, but negatively correlated with wages, or vice-versa. For example, perhaps firms that are less financially constrained have more training provided inside the firm, in programs built by the firm itself. This would mean that these firms are less likely to apply to SENAI training courses, but could still pay higher wages if training is productive. Another explanation for different estimates is the fact that the IV estimates the *local average treatment effect*, measured for the subgroup of individuals for whom the treatment changes their behavior, while the OLS estimates the average effect for the whole population. If the effect of training on wages is heterogeneous across establishments, and lower for the ones that would take up training anyway, this would imply that the IV estimates are higher than the OLS estimates.

The inclusion of firm fixed effects, in columns (3)-(4) and (7)-(8), adds another angle to the story: The OLS estimates suggest returns are no longer positive, and may even be slightly negative. Although this result must be taken with caution, since the IV estimates continue to consistently point to significant positive wage effects, they deliver an interesting conclusion: Firms that train tend to pay high wages overall, but they seem not to discriminate systematically between establishment-occupation groups to which training is provided and establishments-occupation groups that are not trained. This is aligned with the idea that firms tend to set wages nationally so they do not differ considerably from one establishment to the other, even if in some regions workers do not have access to training.

Effect on Wage Dispersion

The literature on training documents that employers have higher incentives to provide training when there is wage compression (Booth and Zoega, 2004), meaning, when the gap in productivity levels between workers is larger than the gap in wages. This is also the case in our theoretical model, where the incentives to train depend positively on the difference between the *"rates of exploitation"*[20] for each skill level. Notice, however, that the existence of wage compression does not imply, either in reality or in our model, that the dispersion of wages in a firm that trains is lower than the dispersion of wages in a firm that does not train. Such a measure depends on the amount of skilled and unskilled workers employed at the firm and on the dispersion of skill levels, as opposed to wage compression.[21] To see this, consider the extreme case of a firm where all workers are paid exactly their marginal productivity, but 99% of workers are unskilled. In this firm, wage compression is low, and inequality is also very low, simply because the distribution of skills is very concentrated on one level. If this firm starts to employ more skilled workers, inequality will increase, but wage compression stays constant as long as the wage policy is unchanged. In fact, our theoretical framework predicts that, everything else constant, the higher the training rate, the higher the difference between the unskilled and the skilled wage. For two firms exactly equal in everything - including labor composition - except in training rates, this means that inequality will be higher in the firm that trains more.

In this section, we focus on the effects of training on wage inequality at the establishment-occupation level. We start by computing the effect of training on general wage dispersion using the variance of log wages as the outcome variable in our baseline regression. The reason for taking the log of wages before computing the variance is two-fold: Firstly, this measure has an advantage when compared to the variance of absolute wages: unlike the former, the variance of log wages is invariant to proportional increases in wages, since the log of a product is equal to the sum of the logs. Additionally, it meets important axioms for inequality measures, such as the axiom of symmetry (a permutation of wages between two workers does not change the inequality index) and the axiom of replication invariance (it remains unchanged if all the incomes are replicated at once in the same income distribution). Besides estimating the impact on the variance, we analyze the impact of training on different ratios between wage percentiles, as is common in the literature on income inequality (for instance, Autor et al. (2008)). Particularly, we estimate the impact on the ratio between the 90^{th} and the 10^{th} percentiles, which is reported to be highly correlated to the Gini coefficient (OECD, 2021), and explore lower and upper tail

[20]This term is used in Manning (2013) to denote the gap between the wage and the marginal product.

[21]For a thorough discussion on this matter, see Tyrowicz and Smyk (2019).

inequality using the P90/P50 and P50/P10 ratios, respectively.

Table 7: Effect of Training on Wage Dispersion, OLS and IV estimates

	Full sample				Excluding Microestablishments			
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Variance of $Ln(wages)$								
$Train_{t-1}$	0.026***	0.070***	0.002***	0.076***	0.018***	0.075***	0.002***	0.085***
	(0.001)	(0.015)	(0.000)	(0.016)	(0.001)	(0.018)	(0.000)	(0.019)
Panel B: P90/P10								
$Train_{t-1}$	0.398***	0.499***	0.122***	0.544***	0.308***	0.485***	0.123***	0.549***
	(0.010)	(0.102)	(0.004)	(0.114)	(0.008)	(0.122)	(0.005)	(0.136)
Panel C: P90/P50								
$Train_{t-1}$	0.109***	0.186***	0.027***	0.205***	0.081***	0.195**	0.027***	0.221**
	(0.003)	(0.046)	(0.002)	(0.051)	(0.003)	(0.057)	(0.002)	(0.063)
Panel D: P50/P10								
$Train_{t-1}$	0.179***	0.434***	0.073***	0.151**	0.139***	0.392***	0.073***	0.266***
	(0.004)	(0.066)	(0.003)	(0.075)	(0.004)	(0.077)	(0.003)	(0.089)
Observations	11,334,794	11,334,794	11,334,794	11,334,794	6,600,055	6,600,055	6,600,055	6,600,055
R-squared	0.138	0.019	0.352	0.005	0.142	0.008	0.284	0.000
Establishment-Occupation controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year × 2-digit sector	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality X Occupation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	Yes	Yes	No	No	Yes	Yes

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Note: This table displays the relationship between training and dispersion of wages at the establishment-occupation level. Establishment-occupation controls include the number of employees and its square, share of women, average worker age and tenure, and share of employees with 9+ and 12+ years of education. For this table, since we are dealing with dispersion of wages, we exclude establishment-occupation units with only one worker, hence the different number of observations. The F-statistic corresponds to the Kleibergen-Paap Wald weak identification test. Standard errors clustered at the municipality × occupation level in parentheses.

Table 7 reports the results of these regressions. As displayed in Panel A, a establishment-occupation group that underwent training has, on average, higher inequality of wages when compared to a unit where no worker underwent training. This is robust to the inclusion of fixed effects, despite a significant drop in the OLS coefficient, which reveals that even two establishments belonging to the same firm will differ in wage dispersion if one has training and the other does not. Turning to the percentile ratios, Panel B confirms that inequality increases among establishment - occupation groups that undergo training, with the IV estimates in our preferred specifications (columns 2 and 6) pointing towards an increase of almost 0.5 in the 90/10 ratio. Furthermore, when we decompose this ratio into two ratios, using the median (in Panels C and D), the results hint at a higher increase in inequality among the workers at the bottom of the wage distribution.

Despite seemingly against the literature reporting a higher prevalence of training when wages are more compressed, our results are compatible with the findings in Bassanini and

Brunello (2003) and Pfeifer (2016) for the aforementioned difference between (de)compression and inequality. Furthermore, these studies measure wage compression at the firm, and at clusters of homogeneous employees, respectively, while we do so at a much finer level: the establishment-occupation group. Our findings may simply mean that groups that are trained are more diverse in terms of skill levels than groups where no worker is trained. Nevertheless, due to the lack of productivity measures in our data, we are not able to disentangle the effects of pure wage (de)compression and skill dispersion. These results may also confirm the theoretical prediction that wages of trained workers will be different from wages of their similar counterparts, even when in the same occupation and establishment (hence the increased dispersion). Hence, further analysis at the worker level is needed to understand what exactly drives these changes in the wage distribution.

Link between Wage Returns and Market Power

As the last set of results presented in this thesis, we now turn our focus to the interplay between training and market power in determining wages. After establishing, in Section 5 that establishments with higher market power are more likely to train, we ask: Do wage returns also differ depending on the market power of the employer? Going back to our theoretical framework in Section 2, the model predicts that bigger firms train more as long as they are able to extract returns from this decision. To test whether we observe this in the data, we estimate a variation of our baseline model 21, where we include a measure of market power in the previous period (P_{ijmt-1}), and a term with the interaction of this measure and enrollment into training,

$$y_{ijmt} = \alpha Train_{ijmt-1} + \delta P_{ijmt-1} + \eta P_{ijmt-1} \times Train_{ijmt-1} + \mathbf{X}'_{ijmt}\beta + \gamma_{mj} + \theta_{st} + \varepsilon_{ijmt} \quad (22)$$

All variables, and indices, have the same meaning as in Equation 21. Notice that while the coefficient associated with P_{ijmt-1}, δ , translates the difference between wages of establishments with different levels of market power that do not train, η, the coefficient associated with the interaction between training and market power, will measure the gap between wages of establishments that train, depending on their market power. Table 8 displays the results of estimating this model with two measures of market power: In Panel A, the logarithm of the raw market share of the firm in the previous year, and in Panel B, a binary variable that takes the value of one if the establishment-occupation group was the largest in its market in the previous period, and zero otherwise.

The results are striking: Across all specifications, the coefficient associated with the interaction of market power and training is consistently negative, meaning, firms with higher market power pay significantly lower returns to training when compared to their similar counterparts with less power in the labor market. If gains in productivity are homogeneous across firms,

Table 8: Effect of Training and Market Power on Average Wage, OLS and IV estimates

	Full sample				**Excluding Microestablishments**			
Dependent Variable: Ln(Average Wage)	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(7)	(8)	(9)	(10)
Panel A: Share								
$Train_{t-1}$	0.081***	-0.030	-0.041***	-0.074	0.059***	0.018	-0.038***	-0.043
	(0.003)	(0.052)	(0.002)	(0.048)	(0.003)	(0.045)	(0.002)	(0.044)
$Ln(share)_{t-1}$	0.057***	0.060***	0.000	0.002	0.031***	0.034***	0.005***	0.007***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
$Ln(share)_{t-1} \times Train_{t-1}$	-0.004**	-0.174***	-0.007***	-0.182***	-0.003*	-0.129***	-0.007***	-0.159***
	(0.002)	(0.043)	(0.001)	(0.038)	(0.002)	(0.036)	(0.001)	(0.034)
Panel B: Top Firm								
$Train_{t-1}$	0.171***	0.213**	-0.026***	0.197**	0.107***	0.160**	-0.019***	0.146***
	(0.003)	(0.084)	(0.003)	(0.076)	(0.003)	(0.063)	(0.002)	(0.053)
Top_{t-1}	0.049***	0.013***	-0.008***	0.012***	0.011***	-0.009**	-0.001	0.021***
	(0.002)	(0.004)	(0.001)	(0.003)	(0.002)	(0.004)	(0.001)	(0.003)
$Top_{t-1} \times Train_{t-1}$	-0.063***	0.159**	-0.062***	-0.320***	-0.036***	0.048	-0.058***	-0.283***
	(0.005)	(0.070)	(0.003)	(0.052)	(0.004)	(0.050)	(0.003)	(0.036)
Observations	19,800,658	19,800,658	19,800,658	19,800,658	9,140,274	9,140,274	9,140,274	9,140,274
R-squared	0.626	0.177	0.826	0.094	0.670	0.223	0.814	0.133
Year X 2-digit sector	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality X Occupation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	Yes	Yes	No	No	Yes	Yes
Fist Stage F-Statistic		87.55		84.47		72.28		72.39

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

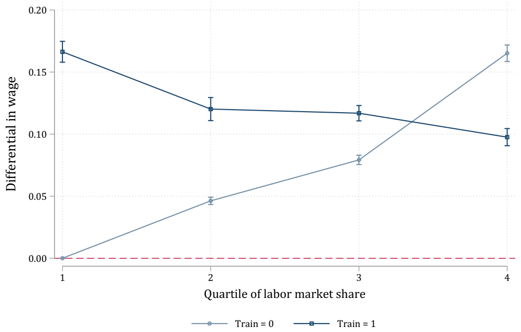
Note: This table displays the estimated effect of training, market power, and their interaction at the establishment-occupation level. Establishments located in municipalities that have no training during the 4-year period are dropped, as well as groups belonging to occupations where no worker is ever trained in the same time window. Establishment-occupation controls include the number of employees and its square, share of women, average worker age and tenure, and share of employees with 9+ and 12+ years of education. The F-statistic corresponds to the Kleibergen-Paap Wald weak identification test. Standard errors clustered at the municipality $\times$ occupation level in parentheses.

then this finding confirms that firms with more market power are better able to extract returns from training their workers, and collect higher benefits from this investment, than less powerful firms. This may well be a large part of the explanation why bigger firms train more.

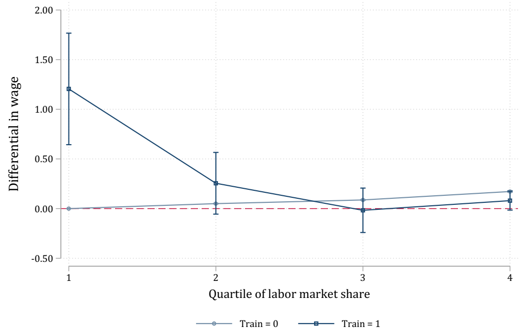
Furthermore, Panel A, in particular, shows that, as soon as market power and the interaction of market power with training are accounted for, the pure returns to training (the coefficient associated with $Train_{t-1}$) drops to zero, or becomes even negative, suggesting that a significant part of the positive effect of training on average wages estimated before may be explained by the fact that training is usually provided by more powerful firms, which tend to pay higher average wages.

Lastly, firms with higher market shares pay more regardless of whether they train or not. This finding is also in line with the predictions in our theoretical model: if productivity differences can, at least partially, explain different training rates and different market shares, then we should expect bigger firms to pay higher average wages. The fact that the coefficient on the market shares drops significantly when we include firm fixed effects also suggests that firm-level differences play a big role in explaining the variability of wages at the establishment-occupation level.

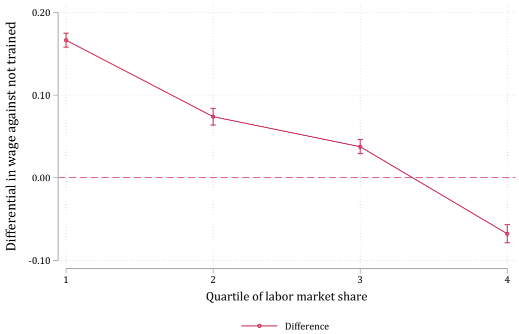
To make quantification of these effects easier, we provide in Figure 5 a visual depiction of the results of estimating model 22 with the quartiles of market share as the measure of market power.



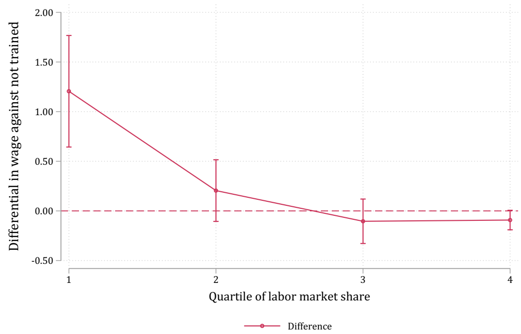
(a) OLS



(b) IV



(c) OLS



(d) IV

Figure 5: Average Wage Returns to Training by quartile

Note: The baseline category are establishments belonging to the first quartile where no training takes place. The values in the vertical axis correspond to percentage differences. Standard errors are clustered municipality $\times$ occupation- level.

The point estimates presented for each of the quartiles correspond to the percentual difference in average wages for establishments in that quartile when compared to similar establishments in the bottom quartile that do not train (our baseline). Graphs (c) and (d) present the difference between the dark blue and the light blue lines of the plots above, and can be interpreted as the percentual net difference in average wages between establishments that train and establishments that do not train in each of the quartiles. The plots show the results of estimating this regression on the full sample. In the Appendix, the same four plots for the sample excluding microestablishments are presented.

The results are aligned with the previous ones: Working in an establishment-occupation group where at least one worker is trained means a higher increase in wage for workers in establishments with low labor market power. In firms with very high market power, which are also the ones where training is more frequent, these programs seem to have the least effect in terms of average wage returns. In fact, among the highest-ranked establishment-occupation units in terms of market power, the difference in average wages between trained and non-trained groups is not statistically different from zero.

7 Robustness Analysis and Additional Results

In this section, we conduct a number of exercises to test the robustness of our results. First, we estimate our main regressions by sector to study the heterogeneity in our results. Second, we construct two alternative instrumental variables, varying the distance threshold considered. Lastly, we answer possible concerns about the validity of our IV by including an extra control.

7.1 Heterogeneity across sectors

As Figure 6 shows, training rates are very heterogeneous across sectors, with most trainees concentrated in a few sectors. To understand whether our main results are heterogeneous across these groups, we run our main regressions by sector, allowing all coefficients to vary between groups. Given that the broadest definition of sector, presented in Table 11, is still quite disaggregated, we further bin sections into three groups, displayed in Table 12.

Table 18, in the Appendix, displays the impact of market power on the likelihood that an establishment-occupation group trains. Although the results are aligned across sectors, they point to a stronger effect of market share on wages in the primary and secondary sectors, than for services. Turning to the effects of training on wages, Table 19, presents the results from estimating our baseline Equation 21 for subsamples by sector. Again, the estimates across sectors are in line with the estimates for the whole sample. Interestingly, the results seem to point to a stronger effect of training on average wages among sectors in *Other Services*. This is surprising given that SENAI training courses are mainly designed for the industry.

Lastly, Table 20 decomposes the effect of training, market power, and the interaction of the two, on wages, by main sector. The effect of market share on wages is generally positive across all sectors and specifications but seems to be stronger in the third group: Other Services. Moreover, it is also in this sector that the differential between average wages of trained and non-trained groups fades faster as market share increases.

7.2 Alternative Instruments

A potential concern with our instrument is the plausibility of the 50 miles threshold chosen. This is particularly relevant given the heterogeneity in municipality areas, as well as road infrastructure, across Brazil. To address this concern, we construct two alternative IV measures that include only courses and workers in a radius of (i) 25 miles and (ii) 75 miles (instead of 50 miles). The First Stage F-statistic is still above 85 for both thresholds, which shows the instrument is not weak. Tables 14 and 16 report the first stage estimates for these alternative instruments. Results for average wage effects using these thresholds do not differ significantly from our baseline results as Tables 15 and 17 report. All other results are quantitatively similar as well.

7.3 Effect of Large Firms

A valid concern regarding our instrument is that the expansion of SENAI courses is, at least in part, demand-driven, and might be particularly influenced by large firms. In an extreme case, if there are only a few large firms in a municipality that are growing over these years, then they may pay higher wages on average and also demand more courses. Therefore, our IV estimates would no longer be valid since we do not meet the exclusion restriction. To address this concern, we estimated all regressions included in the thesis with an extra control for the absolute number of firms in the relevant labor market. None of our results change when we include this control. Particularly, our IV estimates are all not significantly different from the ones where we do not include this extra control.

8 Conclusion

This thesis provides a comprehensive analysis of the relationship between labor market power and firms' decision to engage in job training. Using rich matched employer-employee data for Brazil and linking it to detailed records on training activity from the main provider in Brazil, we find that training is more prevalent in more concentrated markets and that labor market share is an important driver of higher training take-up rates at the firm level. Our estimates suggest that the elasticity of training take-up to market share is close to 3, even when controlling for market concentration. The theoretical prediction that training is more likely when the employer expects to retain the worker for a longer period of time is confirmed in the data, since we find that the separation rate is significantly smaller among establishments with higher market share and in more concentrated markets.

Deploying an instrumental variable approach, we find that training shifts the average of the wage distribution upward, but it also increases the dispersion of wages, especially at the bottom. However, when we take into account the market power of the firm, we show that the positive effect of training on average wages may be partly driven by the correlation between market shares and training rates, since the pure effect of training on wages falls significantly. Lastly, we show that the larger the market power of a firm, the lower the impact of training on average wages.

Many years after the theory shed light on the relevance of market power for the provision of training, our findings confirm the prediction that firms act strategically towards job training, considering the structure of the market in their decision. This evidence suggests that the assumption of perfect competition in labor markets should be questioned. The analysis provided in this thesis can be used to incorporate labor market concentration concerns in the design of effective training programs. By pointing to a positive relationship between market power and firm incentives to train, it draws attention to the need to develop institutions that can foster cooperation and coordination between firms even in highly competitive settings and provide tools for firms to internalize the externalities of human capital accumulation for workers and other employers (for example, through contracts), reducing the risk of poaching and promoting training.

Finally, our findings point to two directions for future research. First, the analysis of the magnitude of the potential distortions caused by the informal sector in this relationship. Specifically, does a larger informal sector reduce the firm incentives to train? Second, the analysis of the impact of training on workers studying long-term effects on wages and contract stability.

References

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. (2017). When should you adjust standard errors for clustering? Technical report, National Bureau of Economic Research.

Acemoglu, D. (1997). Training and innovation in an imperfect labour market. *The Review of Economic Studies*, 64(3):445–464.

Acemoglu, D. and Pischke, J.-S. (1999). Beyond becker: Training in imperfect labour markets. *The economic journal*, 109(453):112–142.

Almeida, R. and Carneiro, P. (2009). The return to firm investments in human capital. *Labour economics*, 16(1):97–106.

Almeida, R. K. and Faria, M. (2014). The wage returns to on-the-job training: evidence from matched employer-employee data. *IZA Journal of Labor & Development*, 3(1):1–33.

Autor, D. H., Katz, L. F., and Kearney, M. S. (2008). Trends in us wage inequality: Revising the revisionists. *The Review of economics and statistics*, 90(2):300–323.

Azar, J., Marinescu, I., and Steinbaum, M. (2022). Labor market concentration. *Journal of Human Resources*, 57(S):S167–S199.

Barron, J. M., Berger, M. C., Black, D. A., et al. (1997). On-the-job training. *Books from Upjohn Press*.

Bartel, A. P. (1995). Training, wage growth, and job performance: Evidence from a company database. *Journal of labor Economics*, 13(3):401–425.

Bassanini, A. and Brunello, G. (2003). Is training more frequent when wage compression is higher? evidence from the european community household panel. *Evidence from the European Community Household Panel (August 2003)*.

Bassanini, A. and Brunello, G. (2008). Is training more frequent when the wage premium is smaller? evidence from the european community household panel. *Labour Economics*, 15(2):272–290.

Bastos, P., Silva, J., and Proença, R. (2016). Exports and job training. *Review of International Economics*, 24(4):737–756.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy*, 70(5, Part 2):9–49.

Benmelech, E., Bergman, N. K., and Kim, H. (2022). Strong employers and weak employees how does employer concentration affect wages? *Journal of Human Resources*, 57(S):S200–S250.

Blanchflower, D., Lynch, L., et al. (1994). Training at work: a comparison of us and british youths. *Training and the private sector*, pages 233–60.

Blyde, J. S., Pires, J. C. L., Chatruc, M. R., and Russell, N. (2019). Training and labor adjustment to trade. Technical report, IDB Working Paper Series.

Bontemps, C., Robin, J.-M., and Van den Berg, G. J. (2000). Equilibrium search with continuous productivity dispersion: Theory and nonparametric estimation. *International Economic Review*, 41(2):305–358.

Booth, A. L. and Snower, D. J. (1996). *Acquiring skills: Market failures, their symptoms and policy responses*. Cambridge University Press.

Booth, A. L. and Zoega, G. (2004). Is wage compression a necessary condition for firm-financed general training? *Oxford Economic Papers*, 56(1):88–97.

Burdett, K. and Mortensen, D. T. (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, pages 257–273.

Card, D. (1999). The causal effect of education on earnings. *Handbook of labor economics*, 3:1801–1863.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5):1127–1160.

Card, D., Cardoso, A. R., Heining, J., and Kline, P. (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics*, 36(S1):S13–S70.

Cassagneau-Francis, O., Gary-Bobo, R., Pernaudet, J., Robin, J.-M., et al. (2020). L'impact de la formation professionnelle en france: une première exploration sur les données defis du céreq. *ÉCHANGES*, page 11.

Conti, G. (2005). Training, productivity and wages in italy. *Labour economics*, 12(4):557–576.

Davis, S. J. and Haltiwanger, J. (1990). Gross job creation and destruction: Microeconomic evidence and macroeconomic implications. *NBER macroeconomics annual*, 5:123–168.

Dearden, L., Reed, H., and Van Reenen, J. (2006). The impact of training on productivity and wages: Evidence from british panel data. *Oxford bulletin of economics and statistics*, 68(4):397–421.

Groes, F., Kircher, P., and Manovskii, I. (2015). The u-shapes of occupational mobility. *The Review of Economic Studies*, 82(2):659–692.

Ibrahim, H., Zin, M. L. M., and Vengdasamy, P. (2020). The influence of individual characteristics, training content and manager support on on-the-job training effectiveness. *The Journal of Asian Finance, Economics, and Business*, 7(11):499–506.

Konings, J. and Vanormelingen, S. (2015). The impact of training on productivity and wages: firm-level evidence. *Review of Economics and Statistics*, 97(2):485–497.

Leitao, M., Montana, J., and Silva, J. (2022). The role of firms in wage inequality dynamics. *Centre for Economic Policy Research.*

Lipsius, B. (2018). Labor market concentration does not explain the falling labor share. *Available at SSRN 3279007.*

Lynch, L. M. (1994). Introduction to" training and the private sector". In *Training and the Private Sector: International Comparisons*, pages 1–24. University of Chicago Press.

Manning, A. (2013). Monopsony in motion. In *Monopsony in Motion*. Princeton University Press.

Muehlemann, S. and Wolter, S. C. (2011). Firm-sponsored training and poaching externalities in regional labor markets. *Regional Science and Urban Economics*, 41(6):560–570.

Muendler, M.-A., Poole, J., Ramey, G., and Wajnberg, T. (2004). Job concordances for brazil: Mapping the classificaçao brasileira de ocupaçoes (cbo) to the international standard classification of occupations (isco-88). *University of California, San Diego, unpublished manuscript.*

OECD (2021). *Old-age income inequality.*

Pfeifer, C. (2016). Intra-firm wage compression and coverage of training costs: Evidence from linked employer-employee data. *ILR Review*, 69(2):435–454.

Rinz, K. (2022). Labor market concentration, earnings, and inequality. *Journal of Human Resources*, 57(S):S251–S283.

Stock, J. H., Wright, J. H., and Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4):518–529.

Tyrowicz, J. and Smyk, M. (2019). Wage inequality and structural change. *Social Indicators Research*, 141(2):503–538.

Webber, D. A. (2015). Firm market power and the earnings distribution. *Labour Economics*, 35:123–134.

Appendix A: Tables

Table 9: Descriptive Statistics for establishment-occupation groups (excluding microestablishments)

	Trains at least once		Never Trains		
	Mean	S.D.	Mean	S.D.	Difference
Number of employees	60.469	779.243	7.419	34.612	-53.050***
Average Real Wage	125,074.073	121,698.256	111,739.335	153,355.917	-13,334.738***
Average employee age	34.041	6.813	35.119	9.054	1.078***
Average employee tenure	3.826	3.931	3.569	4.354	-0.257***
Share of women employees	0.251	0.313	0.422	0.422	0.171***
Share of employees with 9+ years of education	0.882	0.211	0.891	0.257	0.010***
Share of employees with 12+ years of education	0.671	0.335	0.689	0.397	0.017***
Share of employees with a university degree	0.122	0.261	0.169	0.340	0.047***
Number of trained employees	1.349	9.458	0.000	0.000	-1.349***
Observations	1,007,888		9,572,965		

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: This table presents the average of some variables of interest for establishment-occupation where at least one worker is trained in the period 2009-2012 and for establishment-occupation groups where no worker is trained in the same period. The sample excludes microestablishments. Real wages are obtained from annual nominal wages which we deflate using a consumer price index for Brazil.

Table 10: Sample Size

	Number of single ID's	
Level	Full sample	Excluding Microest.
Firm	2,750,437	563,157
Establishment	3,253,258	704,395
Establishment-Occupation	9,769,429	4,201,926

The table displays the number of single units for each aggregation level included in our analysis. It already excludes units in labor markets with less than 10 workers, as well as singleton firms (which would be dropped once we include firm fixed effects).

Table 11: National Classification of Economic Activities (CNAE)

Section	Divisions	Description (author's own translation)
A	01 .. 03	Agriculture, livestock, forestry, fisheries and aquaculture
B	05 .. 09	Extractive industries
C	10 .. 33	Manufacturing industries
D	35 .. 35	Electricity and gas
E	36 .. 39	Water, sewage, waste management activities and decontamination
F	41 .. 43	Construction
G	45 .. 47	Trade; Repair of motor vehicles and motorcycles
H	49 .. 53	Transportation, storage, and mail
I	55 .. 56	Accommodation and food
J	58 .. 63	Information and communication
K	64 .. 66	Financial, insurance, and related services
L	68 .. 68	Real estate activities
M	69 .. 75	Professional, scientific and technical activities
N	77 .. 82	Administrative activities and complementary services
O	84 .. 84	Public administration, defense, and social security
P	85 .. 85	Education
Q	86 .. 88	Human health and social services
R	90 .. 93	Arts, culture, sport, and recreation
S	94 .. 96	Other service activities
T	97 .. 97	Domestic services
U	99 .. 99	International organizations and other extraterritorial institutions

The table displays the most aggregated definition of sector using the Brazilian CNAE classification of economic activity. The definition in the first collumn *(Section)* is the one used in the labor market definition in the robustness test, as is also the one used to study heterogeneity of effects in section XXX. The second collumn *(Division)* presents the correspondent two-digit classifications for each section. This is the level used for the sector fixed effects, used in all specifications throughout the thesis.

Table 12: Aggregation of Sectors

Divisions	Aggregated Sector
01 .. 43	Primary and Secondary Sectors
45 .. 63	Trade, Transportation, Accomodation and Communication
64 .. 99	Other services

The table displays our aggregation of sectors in 3 broad categories to study heterogeneity of returns to training and their interplay with market power across sectors. The first column *(Division)* presents the correspondent two-digit classifications for each aggregation.

Table 13: National classification of Occupations (CBO)

Code	Description
1	Members of the armed forces
2	Military police officers
3	Military firefighters
11	Senior members and leaders of public authorities
12	Directors of companies and organizations (except public interest)
13	Directors and managers in health services, education, or services...
14	Managers
20	Researchers and political professionals
21	Exact, physical and engineering science professionals
...	
81	Workers in continuous process industries and other industries
82	Steel plant and building materials workers
83	Pulp and paper machinery and plant workers
84	Food, beverage and tobacco production workers
86	Production, gathering, treatment and distribution operators
87	Operators of other industrial facilities
91	Mechanical repair and maintenance service workers
95	Polymantine workers
99	Other maintenance, repair and maintenance workers

The table displays the description of some occupations according to the Brazilian classification of occupations (*Classificação Brasileira de Ocupações - CBO*), used in this thesis for the definition of labor market.

Table 14: First-stage regression: Alternative IV (25 miles)

	Full sample		**Excluding Microest.**	
Dependent variable: $Train_t$	(1)	(2)	(3)	(4)
Courses 25 miles / number of workers	0.147***	0.145***	0.184***	0.178***
(excluding own)	(0.011)	(0.011)	(0.015)	(0.014)
Observations	19,800,658	19,800,658	9,140,274	9,140,274
Establishment-Occupation Controls	Yes	Yes	Yes	Yes
Year $\times$ 2-digit sector	Yes	Yes	Yes	Yes
Municipality X Occupation	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Note: This table displays the first-stage estimates for an alternative to our preferred IV. Establishment-occupation controls include the number of employees and its square, share of women, average worker age and tenure, and share of employees with 9+ and 12+ years of education. Standard errors clustered at the municipality $\times$ occupation level in parentheses.

Table 15: Average Wage Returns: Alternative IV (25 miles), OLS and IV estimates

	Full sample		**Excluding Microest.**	
Dependent variable: Ln(Average Wage)	(1)	(2)	(3)	(4)
$Train_{t-1}$	0.184**	0.159**	0.147**	0.122**
	(0.074)	(0.067)	(0.060)	(0.050)
Observations	19,800,658	19,800,658	9,140,274	9,140,274
Establishment-Occupation Controls	Yes	Yes	Yes	Yes
Year $\times$ 2-digit sector	Yes	Yes	Yes	Yes
Municipality X Occupation	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes
First Stage F-Statistic	96.23	101.6	91.23	95.76

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Note: This table displays the relationship between training and average wages at the establishment-occupation level, using an alternative IV specification with a threshold of 25 miles, instead of 50. Establishment-occupation controls include the number of employees and its square, share of women, average worker age and tenure, and share of employees with 9+ and 12+ years of education. The F-statistic corresponds to the Kleibergen-Paap Wald weak identification test. Standard errors clustered at the municipality $\times$ occupation level in parentheses.

Table 16: First-stage regression: Alternative IV (75 miles)

	Full sample		**Excluding Microest.**	
Dependent variable: $Train_t$	(1)	(2)	(3)	(4)
Courses 25 miles / number of workers	0.197***	0.190***	0.248***	0.237***
(excluding own)	(0.014)	(0.014)	(0.019)	(0.018)
Observations	19,800,658	19,800,658	9,140,274	9,140,274
Establishment-Occupation Controls	Yes	Yes	Yes	Yes
Year $\times$ 2-digit sector	Yes	Yes	Yes	Yes
Municipality X Occupation	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Note: This table displays the first-stage estimates for an alternative to our preferred IV. Establishment-occupation controls include the number of employees and its square, share of women, average worker age and tenure, and share of employees with 9+ and 12+ years of education. Standard errors clustered at the municipality $\times$ occupation level in parentheses.

Table 17: Average Wage Returns: Alternative IV (75 miles), OLS and IV estimates

	Full sample		Excluding Microest.	
Dependent variable: Ln(Average Wage)	(1)	(2)	(3)	(4)
$Train_{t-1}$	0.237***	0.187**	0.170***	0.128***
	(0.077)	(0.073)	(0.058)	(0.049)
Observations	19,800,658	19,800,658	9,140,274	9,140,274
Establishment-Occupation Controls	Yes	Yes	Yes	Yes
Year $\times$ 2-digit sector	Yes	Yes	Yes	Yes
Municipality X Occupation	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes
First Stage F-Statistic	88.90	90.64	86.56	88.29

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Note: This table displays the relationship between training and average wages at the establishment-occupation level, using an alternative IV specification with a threshold of 75 miles, instead of 50.Establishment-occupation controls include the number of employees and its square, share of women, average worker age and tenure, and share of employees with 9+ and 12+ years of education. The F-statistic corresponds to the Kleibergen-Paap Wald weak identification test. Standard errors clustered at the municipality $\times$ occupation level in parentheses.

Table 18: Effect of Market Share on Training Rates by sector, OLS and IV estimates

	Full sample		Excluding Microest.	
Dependent variable: $Train_t$	(1)	(2)	(3)	(4)
Panel A: Primary and Secondary Sectors				
$Ln(share)$	0.040***	0.034***	0.056***	0.047***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	4,163,963	4,153,545	1,947,417	1,947,385
R-squared	0.148	0.252	0.178	0.258
Panel B: Trade, Transportation and Communication				
$Ln(share)$	0.012***	0.011***	0.017***	0.016***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	13,976,325	13,955,387	5,490,555	5,490,143
R-squared	0.073	0.192	0.095	0.195
Panel C: Other Services				
$Ln(share)$	0.021***	0.021***	0.027***	0.025***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	6,922,684	6,910,470	3,142,843	3,142,618
R-squared	0.075	0.203	0.093	0.204
Establishment-Occupation Controls	Yes	Yes	Yes	Yes
Year $\times$ 2-digit sector	Yes	Yes	Yes	Yes
Municipality X Occupation	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Note: This table displays the outcomes of estimating a regression whose dependent variable is equal to one if the firm trains at least one worker and zero otherwise with fixed effects estimation in three subsamples by sector as described in Table 12. Collumns (1)-(2) show the effect of market power and market concentration on the probability of training using all establishment-occupation groups in labor markets with at least 10 employees. In columns (3)-(4), groups that belong to microestablishments are excluded from the sample. Standard errors clustered at the municipality $\times$ occupation level in parentheses.

Table 19: Effect of Training on Average Wages by sector, OLS and IV estimates

	Full sample				**Excluding Microestablishments**			
Dependent Variable: Ln(Average Wage)	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Primary and Secondary Sectors								
$Train_{t-1}$	0.146***	0.105*	-0.033***	0.146***	0.090***	0.103**	-0.027***	0.112***
	(0.002)	(0.060)	(0.002)	(0.051)	(0.002)	(0.052)	(0.002)	(0.042)
Observations	3,274,472	3,274,472	3,268,446	3,268,446	1,667,840	1,667,840	1,667,685	1,667,685
R-squared	0.688	0.250	0.845	0.201	0.737	0.295	0.842	0.238
Panel B: Trade, Transportation, Communication								
$Train_{t-1}$	0.141***	0.168	-0.011***	0.065	0.096***	0.188**	-0.007**	0.093
	(0.004)	(0.109)	(0.003)	(0.096)	(0.003)	(0.085)	(0.003)	(0.072)
Observations	10,673,250	10,673,250	10,657,861	10,657,861	4,629,769	4,629,769	4,628,898	4,628,898
R-squared	0.565	0.124	0.799	0.122	0.617	0.173	0.783	0.155
Panel C: Other services								
$Train_{t-1}$	0.152***	0.741***	-0.014***	0.701***	0.101***	0.313**	-0.009***	0.299**
	(0.006)	(0.214)	(0.003)	(0.211)	(0.006)	(0.140)	(0.003)	(0.127)
Observations	5,595,035	5,595,035	5,587,698	5,587,698	2,716,035	2,716,035	2,715,577	2,715,577
R-squared	0.651	0.202	0.851	0.082	0.685	0.265	0.836	0.166
Establishment-Occupation Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year $\times$ 2-digit sector	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality X Occupation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	Yes	Yes	No	No	Yes	Yes

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: This table displays the relationship between training and average wages at the establishment-occupation level by aggregated sector as described in Table 12. Establishment-occupation controls include the number of employees and its square, share of women, average worker age and tenure, and share of employees with 9+ and 12+ years of education. The F-statistic corresponds to the Kleibergen-Paap Wald weak identification test. Standard errors clustered at the municipality $\times$ occupation level in parentheses.

Table 20: Effect of Training and Market Power on Average Wage by sector, OLS and IV estimates

	Full sample				Excluding Microestablishments			
Dependent Variable: Ln(Average Wage)	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(7)	(8)	(9)	(10)
Panel A: Primary and Secondary Sectors								
$Train_{t-1}$	0.080***	0.054*	-0.025***	0.029	0.054***	0.097**	-0.021***	0.071
	(0.003)	(0.033)	(0.002)	(0.042)	(0.002)	(0.039)	(0.002)	(0.043)
$Ln(share)_{t-1}$	0.057***	0.059***	-0.014***	-0.014***	0.035***	0.034***	-0.010***	-0.010***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)
$Ln(share)_{t-1} \times Train_{t-1}$	-0.006***	-0.087**	-0.003***	-0.131***	-0.004***	-0.043	-0.003***	-0.117***
	(0.001)	(0.039)	(0.001)	(0.030)	(0.001)	(0.039)	(0.001)	(0.033)
Observations	3,274,472	3,274,472	3,268,446	3,268,446	1,667,840	1,667,840	1,667,685	1,667,685
R-squared	0.694	0.254	0.845	0.151	0.739	0.296	0.842	0.183
Panel B: Trade, Transportation and Communication								
$Train_{t-1}$	0.099***	-0.186**	-0.016***	-0.068	0.074***	-0.060	-0.017***	-0.023
	(0.004)	(0.089)	(0.002)	(0.045)	(0.003)	(0.067)	(0.003)	(0.044)
$Ln(share)_{t-1}$	0.055***	0.055***	0.003**	0.002*	0.028***	0.027***	0.008***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
$Ln(share)_{t-1} \times Train_{t-1}$	0.000	-0.186***	-0.002	-0.068	-0.001	-0.142***	-0.003	-0.065
	(0.002)	(0.068)	(0.001)	(0.050)	(0.002)	(0.049)	(0.002)	(0.040)
Observations	10,673,250	10,673,250	10,657,861	10,657,861	4,629,769	4,629,769	4,628,898	4,628,898
R-squared	0.573	0.117	0.799	0.117	0.619	0.157	0.783	0.149
Panel C: Other Services								
$Train_{t-1}$	0.010	-0.394	-0.031***	-0.091	0.022***	-0.245	-0.029***	-0.141
	(0.006)	(0.275)	(0.003)	(0.209)	(0.006)	(0.187)	(0.003)	(0.158)
$Ln(share)_{t-1}$	0.065***	0.065***	0.010***	0.008	0.039***	0.043***	0.013***	0.013***
	(0.002)	(0.006)	(0.001)	(0.005)	(0.002)	(0.006)	(0.001)	(0.005)
$Ln(share)_{t-1} \times Train_{t-1}$	-0.017***	-0.791***	-0.003***	-0.630***	-0.012***	-0.516***	-0.002***	-0.472***
	(0.003)	(0.223)	(0.001)	(0.199)	(0.003)	(0.190)	(0.001)	(0.182)
Observations	5,595,035	5,595,035	5,587,698	5,587,698	2,716,035	2,716,035	2,715,577	2,715,577
R-squared	0.660	-0.150	0.851	-0.485	0.689	0.037	0.836	-0.240
Year X 2-digit sector	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality X Occupation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	Yes	Yes	No	No	Yes	Yes

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: This table displays the estimated effect of training, market power, and their interaction at the establishment-occupation level by sector as described in Table 12. Establishment-occupation controls include the number of employees and its square, share of women, average worker age and tenure, and share of employees with 9+ and 12+ years of education. The F-statistic corresponds to the Kleibergen-Paap Wald weak identification test. Standard errors clustered at the municipality $\times$ occupation level in parentheses.

Appendix B: Figures

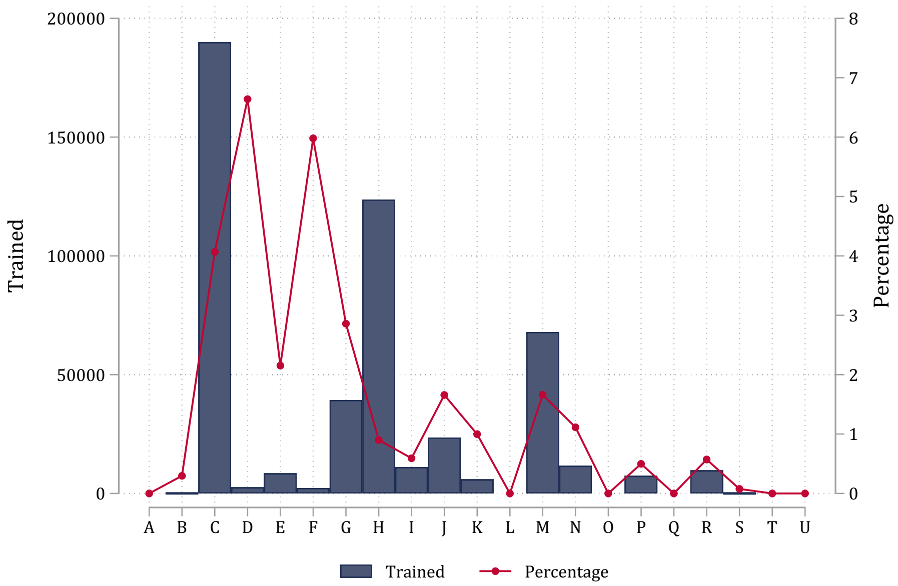


Figure 6: Establishment-occupation groups trained by Sector

Note: The figure displays the number and the percentage of establishment-occupation groups trained by sector of activity, averaging over the 2009-2012 period. For the correspondence between the letters in the horizontal axis (sections) and the sectors, see Table 11.

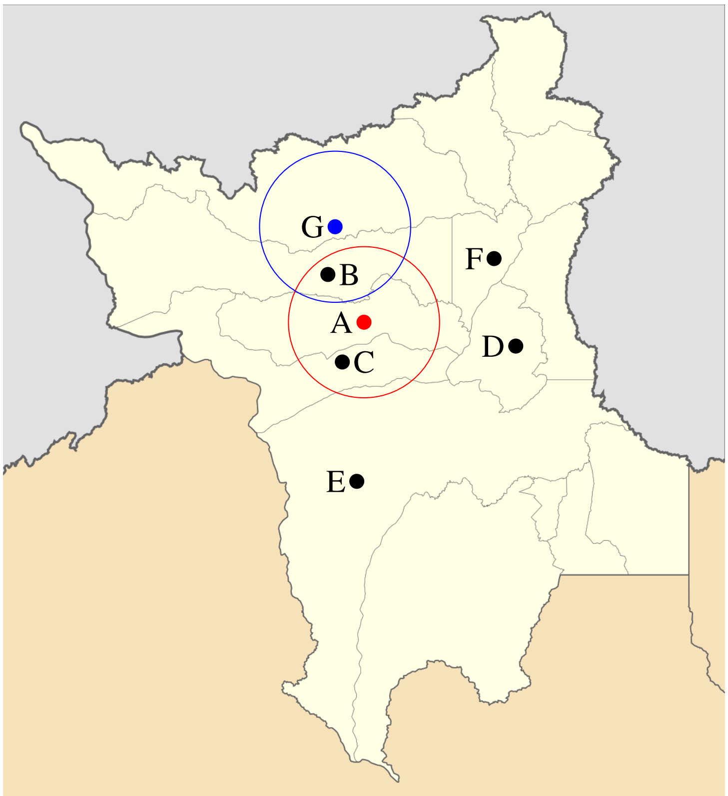


Figure 7: Clarification of the *50 miles radius* definition used

Note: The map represents the municipalities that would be included in the calculation of the number of courses and number of workers within 50 miles of the establishments in municipality A (red circle) and in municipality G (blue circle). Let us consider that the circumference drawn has a radius equal to 50 miles. In this fictitious example (where the centers do not necessarily correspond to the real coordinates), all workers and courses in municipalities A, B and C would be included in the instrument of municipality A, but we would exclude B,E,F and D. For municipality G, only workers and courses from municipalities G and B would be included, and we would exclude those in A,C,D,E and F.

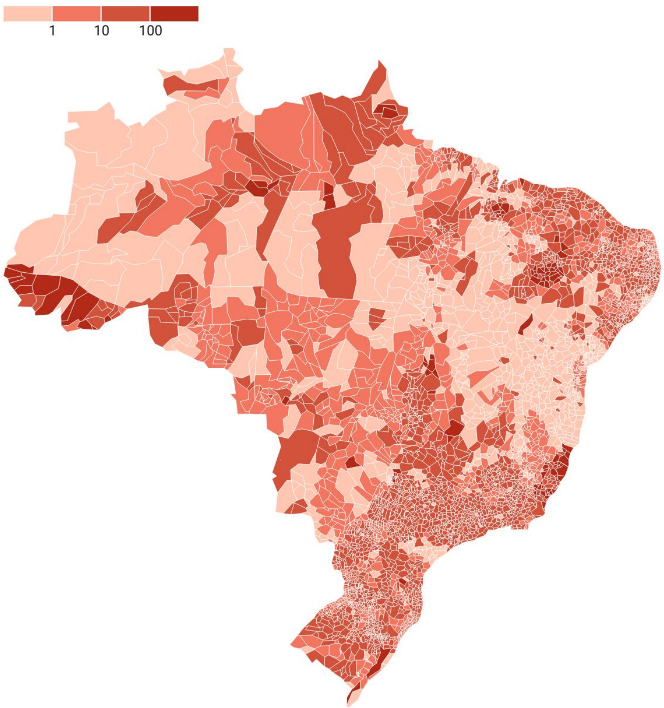
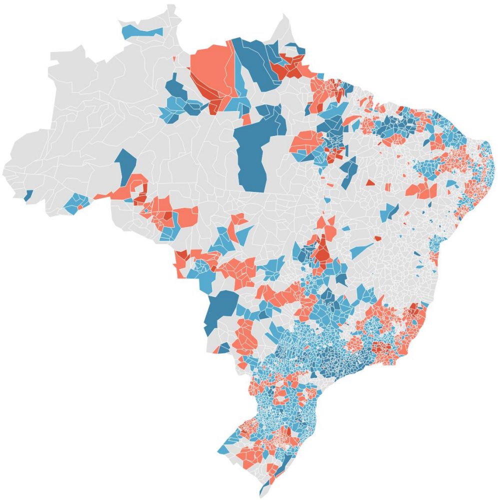
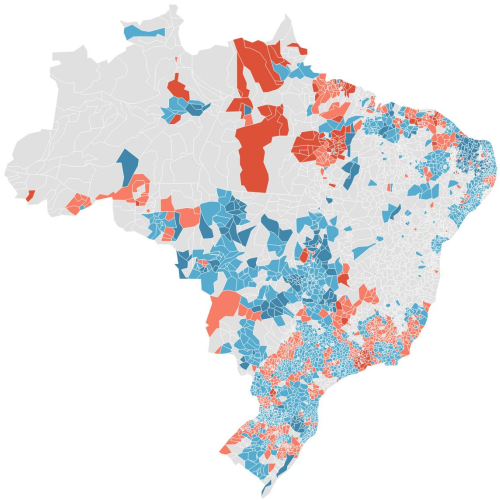


Figure 8: SENAI courses per 1000 workers

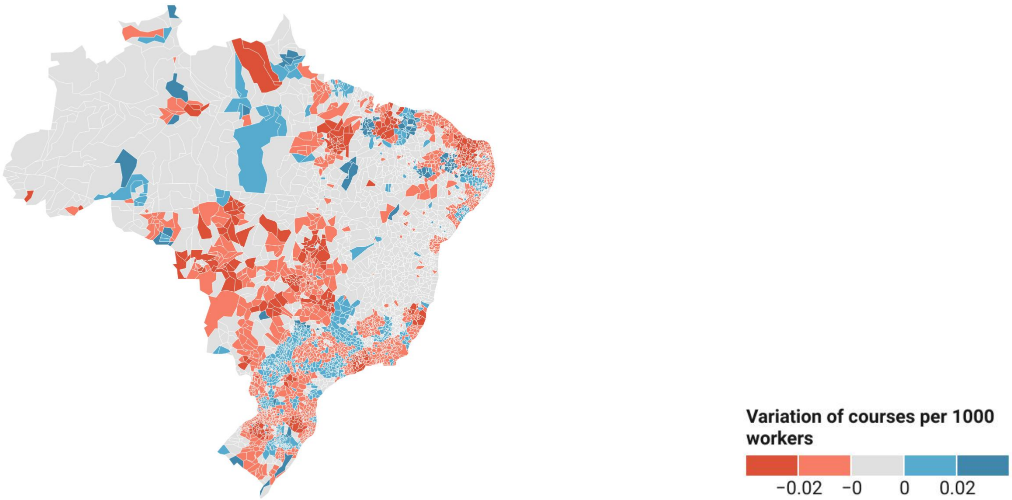
Note: The map presents the base for computing our preferred IV. It depicts the average number of SENAI technical upgrading courses offered within 50 miles of the municipality per 1000 workers in the same area (for readability, we use the measure per 1000 workers instead of per worker) for the period 2009-2013.



(a) 2009-2010



(b) 2010-2011

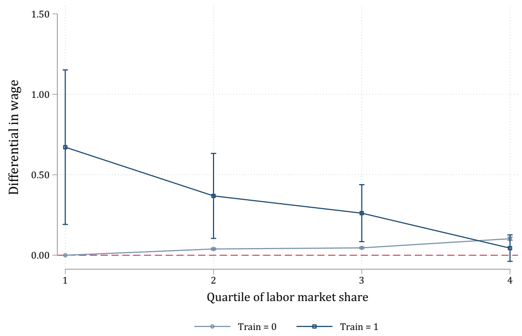


(c) 2011-2012

Figure 9: Yearly variation of the Instrumental Variable
Occupation 78 - Cross-Functional workers

Note: The graphs display the yearly variation in the number of courses per 1000 workers by municipality, for occupation 78 (cross-functional workers).

(a) OLS



(b) IV

(c) OLS

(d) IV

Figure 10: Average Wage Returns to Training by quartile

Note: Note: The baseline category are establishments belonging to the first quartile where no training takes place. The values in the vertical axis correspond to percentage differences. The sample excludes microestablishments. Standard errors are clustered at the municipality $\times$ occupation- level.

Appendix C: Theoretical Model

The derivation of the theoretical model, presented in this section, follows closely the approach to the Burdett and Mortensen's model proposed in Manning (2013). However, we differ from this model in two key aspects: First, we generalize this framework to the case where there is a distribution of productivity levels in the economy. Secondly, we assume a specific technology for training where the increase in productivity arising from it is a fixed proportion of the initial worker productivity at the firm.

Assumptions

We start by making a set of assumptions about the labor market:

1. There is a unitary mass of workers, all of them equally productive at the outset. All workers are unskilled ($T = 0$) when they enter the labor force, but they can become skilled ($T = 1$) if they are trained by the firm [22]. These workers attach an equal value, b, to leisure.

2. There is a unitary mass of employers (firms), whose technology has constant returns to scale. The production functions of these firms differ in A_i - which is a shock to productivity. There is a continuous distribution of productivity levels in the economy. The production function of firm i can be written as:

$$Y_i = p_{i0}N_{i0} + p_{i1}N_{i1} \tag{23}$$

   where N_{i0} is the amount of unskilled labor and N_{i1} is the amount of skilled labor employed at the firm. Denote the productivity of an unskilled individual working in firm i, p_{i0} and the productivity of a skilled individual working in the same firm p_{i1}.

3. Firms can offer two types of contracts: an unskilled and a skilled wage. Firms can transform an unskilled worker into a skilled one by training the worker in the firm (and they can charge directly for training[23]). $\chi_i = \frac{T}{N_{i0}}$ corresponds to the training intensity in the firm. Meaning, it is the rate at which unskilled workers are converted into skilled workers.

4. Both employed and non-employed workers (regardless of their skill level) receive job offers at rate λ (which are assumed to be evenly distributed among the mass of firms and among all individuals), leave the labor force at rate δ_r and move from employment to unemployment at the job-destruction rate δ_u. For future use, we define $\delta = \delta_r + \delta_u$.

[22] This model assumes implicitly that the firm is the only provider of training in the economy

[23] As shown later, charging directly for training or not is irrelevant for the profits of the firm

Workers accept any job offer if the value of the job offer is above their current value or the value of being unemployed.

5. Denote by $s_{iT}(w_{iT})$ and $R_{iT}(w_{iT})$, $T = 0, 1$ the separation and recruitment rates for unskilled and skilled workers, which, in a monopolistic labor market, depend on the wage paid (and so may differ between firms as long as firms pay different wages).

Worker

Workers maximize their lifetime utility. The value of being an unskilled worker in firm i, V_{i0}, is:

$$\delta_r V_{i0} = w_{i0} - \delta_u[V_{i0} - V_0^u] + \chi_i(V_{i1} - V_{i0}) + \lambda \int_{V_{i0}}^{1} [V - V_{i0}] dF_0(V) \tag{24}$$

where w_{i0} is the wage, $F_0(V)$ is the distribution of values on offer to the unskilled in the external labor market and V_{i1} is the value of being a skilled worker in the firm. Notice that the value of being an unskilled worker incorporates not only the wage, but also the probability of being unemployed, and the possibility of becoming a skilled worker (which depends on the training intensity at the firm, χ) or receiving a better offer.

The value of being trained, for a worker, is given by ($V_{i1} - V_{i0}$). Therefore, the worker is willing to pay exactly this amount to get training. Despite rarely observed, we will assume, in this model, that there is a direct charge for training. We provide a detailed proof at the end of this section showing that this assumption is worth nothing to the employer, who makes the same profits. However, it simplifies significantly our problem: Since the worker pays exactly the total amount of future benefits to undergo training, it is only the unskilled wage that affects the value of a job to an unskilled worker (and not the probability of becoming skilled). This assumption implies the following:

Proposition 5 *The relationship between the wage of unskilled workers when they do not pay a direct cost for training, $\bar{w_0}$, and the wage when they do pay a direct cost for training, w_0, is given by:*

$$\bar{w_0} = w_0 - \chi(V_1 - V_0) \tag{25}$$

Here we omit the firm subscript for simplicity of notation. The value of a job for an unskilled worker is then given by

$$\delta_r V_{i0} = w_{i0} - \delta_u[V_{i0} - V_0^u] + \lambda \int_{w_0} [V(w_0) - V(w_{i0})] dF_0(w_0) \tag{26}$$

For skilled workers, we have:

$$\delta_r V_{i1} = w_1 - \delta_u[V_{i1} - V_1^u] + \lambda \int_{w_{i1}} [V(w_1) - V(w_{i1})] dF_1(w_1) \tag{27}$$

At this point, it is important to clarify that the firm training rates do not affect the worker decision to accept an offer.

Proposition 6 *A worker, regarding of its skill level, only cares about the wage of an offer, when deciding whether to take the offer or not.*

Proposition 6 may raise some questions at first. But there are two key assumptions worth taking into account: First, workers are indifferent between taking training or not, since we are assuming that they pay a direct cost for training, which is exactly equal to the difference of values between being skilled and unskilled. Secondly, accepting an offer at any point does not prevent the worker from accepting a better offer in the future. The arrival rate of offers is the same (λ) regarding of the state of the worker (skilled, unskilled, employed or unemployed) so that the value of two distinct offers only differs in the current wage.

Steady-state conditions

In the steady-state, each firm i employs a constant number of skilled and unskilled workers. Therefore:

$$s_{i0}\{w_{i0}; F_0\} N_{i0} = R_{i0}(w_{i0}; F_0) - T \tag{28}$$

$$s_{i1}\{w_{i1}; F_1\} N_{i1} = R_{i1}(w_{i1}; F_1) + T \tag{29}$$

Rearranging 28 and 29, one can write the steady-state level of employment in firm i as:

$$N_{i0} = \frac{R_{i0}(w_{i0}; F_0) - T}{s_{i0}(w_{i0}; F_0)} \tag{30}$$

$$N_{i1} = \frac{R_{i1}(w_{i1}; F_1) + T}{s_{i1}(w_{i1}; F_1)} \tag{31}$$

Note that, although firms differ in productivity levels, the recruitment and separation rates depend only on the wage offered by a given firm and the general distribution of wages in the economy, F_0 and F_1.

Firm's profits

The profits of the firm can be written as:

$$\pi_i = (p_{i0} - w_{i0})N_{i0} - (p_{i1} - w_{i1})N_{i1} - [c - (V_{i1} - V_{i0})]T_i \tag{32}$$

Given 30 and 31, one can write profits as:

$$\pi_i = (p_{i1} - w_{i1})\frac{R_{i1}(w_{i1}; F_1) + T_i}{s_{i1}(w_{i1}; F_1)} + (p_{i0} - w_{i0})\frac{R_{i0}(w_{i0}; F_0) - T_i}{s_0(w_{i0}; F_0)} - [c - (V_1 - V_0)]T_i \tag{33}$$

Firms maximize profits choosing $\{w_0, w_1, T\}$ The first order condition for unskilled wage can be written as:

$$(p_{i0} - w_{i0})\frac{N_{i0}}{w_{i0}} - N_{i0} - \frac{V_{i0}}{w_{i0}}T_i = 0 \tag{34}$$

The first order condition for skilled wage can be written as:

$$(p_{i1} - w_{i1})\frac{N_{i1}}{w_{i1}} - N_{i1} + \frac{V_{i1}}{w_{i1}}T_i = 0 \tag{35}$$

We can simplify this conditions so that they yield:

$$w_{i0} = \frac{\varepsilon_{i0}^R - \varepsilon_{i0}^S(1 - \theta_{i0})}{1 + \varepsilon_{i0}^R - \varepsilon_{i0}^S(1 - \theta_{i0})}p_{i0} \tag{36}$$

where $\theta_0 = \frac{T}{R_0}$

$$w_{i1} = \frac{\varepsilon_{i1}^R - \varepsilon_{i1}^S(1 + \theta_{i1})}{1 + \varepsilon_{i1}^R - \varepsilon_{i1}^S(1 + \theta_{i1})}p_{i1} \tag{37}$$

where $\theta_{i1} = \frac{T_i}{R_{i1}} = \theta_{i0}\frac{R_{i0}}{R_{i1}}$. ε_T^R and ε_T^S represent, respectively, the elasticity of recruitment and the elasticity of separation with respect to the wage rate for each of the skill levels $T = 0, 1$.

The first order condition for the level of training can be written as:

$$[\frac{p_{i1} - w_{i1}}{s_{i1}} + V_{i1}] - [\frac{p_{i0} - w_{i0}}{s_{i0}} + V_{i0}] - c = 0 \tag{38}$$

This first term in this condition represents the marginal benefit of training one more worker for both worker and firm: the difference between the productivity increase and the wage increase accounting for the separation rate plus the stream of future benefits for the worker from being skilled. The second and third term represent the cost from training one more worker: the loss of an unskilled workers plus the direct cost from training.

Proof of Proposition 5

(This proof was directly taken from Manning 2003)

For simplicity, we omit the firm subscript i. Consider the case where workers are not charged directly for their training. Then, if the firm trains unskilled workers at rate χ, the value of being unskilled at the firm is given by:

$$\delta_r V_0 = \bar{w_0} - \delta_u[V_0 - V_0^u] + \chi(V_1 - V_0) + \lambda \int_{V_0}^{1} [V - V_0] dF_0(V) \tag{39}$$

where $F_0(V)$ is the distribution of values on offer to the unskilled in the external labor market (offers differ not only on the wage but also on the probability of being trained at that firm) and V_1 is the value of being a skilled worker in the firm. Similarly, for skilled workers, we have:

$$\delta_r V_1 = \bar{w_1} - \delta_u[V_1 - V_1^u] + \lambda \int_{V_1}^{1} [V - V_1] dF_1(V) \tag{40}$$

Notice that this value no longer depends on the training intensity at the firm. Now, profits can be written as:

$$\bar{\pi} = (p_1 - \bar{w_1})N_1 + (p_0 - \bar{w_0})N_0 - cT \tag{41}$$

Let us suppose now that the employer changes from a contract in which workers are not charged for training to one in which they are. The amount that can be charged is the gain to them from becoming skilled, which is given by $(V_1 - V_0)$. Furthermore, consider that, in making this change, the employer does not want to change the value of being a skilled or unskilled worker in the firm. As unskilled workers no longer gain from becoming skilled, Equation 39 implies that the wage now paid to these workers, w_0 must be:

$$\bar{w_0} = w_0 - \chi(V_1 - V_0) \tag{42}$$

while the skilled wage remains unchanged. With this change, and with the fact that it now costs employers $c - (V_1 - V_0)$ whenever workers are trained, employers will make the following level of profits in the new contract:

$$\begin{aligned}
\pi =& (p_1 - w_1)N_1 + (p_0 - w_0)N_0 - [c - (V_1 - V_0)]T \\
=& (p_1 - \bar{w_1})N_1 + (p_0 - \bar{w_0})N_0 - \chi(V_1 - V_0)N_0 - [c - (V_1 - V_0)]T \\
=& \bar{\pi}
\end{aligned} \tag{43}$$