

THE EFFECTS OF MANDATING TRAINING IN FIRMS: THEORY AND EVIDENCE FROM THE COLOMBIAN APPRENTICESHIP PROGRAM *

SANTIAGO CAICEDO, MIGUEL ESPINOSA AND ARTHUR SEIBOLD

February 15, 2019

Preliminary and Incomplete (Please do not circulate)

Abstract

We study the effect of apprenticeship programs on firms and welfare, using novel administrative data on the universe Colombian manufacturing firms with at least 10 workers, and a unique reform to apprenticeship regulation. The reform simultaneously establishes apprentice quotas that vary discontinuously in firm size and lowers apprentices' wages. We begin by documenting that the policy is successful in increasing the number of trained apprentices more than threefold. However, the reform also induces significant firm size distortions driven by heterogeneous firm responses. In sectors with high skill requirements, firms avoided hiring apprentices decreasing their size and bunching just below the regulation thresholds. In contrast, firms in low-skilled sectors, increase their size and bunch just above the regulation thresholds in order to be able to hire more apprentices. As a consequence, the regulation results in most apprentices being trained in low-skilled sectors. We develop a simple theoretical model featuring heterogeneous training costs across sectors in order to rationalize and quantify these empirical findings. The key insight of the model is firms that train apprentices incur in an opportunity cost of spending time teaching and not producing. As training in high-skill sectors takes longer than in low-skill sectors, firms in high skilled sectors will avoid apprentices while firms in low-skill sectors try to get as many as possible. Finally, we use the model to analyze the welfare consequences of the regulation and study counterfactual policies.

Key words: Training, Apprentices, Firm-size distortions

JEL Codes: E24, J21, J24.

*Santiago Caicedo, Universidad de los Andes, scaicedosoler@uchicago.edu, s-caice1@uniandes.edu.co; Miguel Espinosa, Universitat Pompeu Fabra, miguel.espinosa@upf.edu ; Arthur Seibold, University of Mannheim, seibold@uni-mannheim.de . We would like to thank Ufuk Akcigit, Benito Arrunada, Leonardo Bursztyn, Jonathan Dingel, Rosa Ferrer, Luis Garicano, Alessandro Gavazza, Maria Guadalupe, Ali Hortacsu, Camille Landais, Bob Lucas, Sandra McNally, Joan Monras, Massimo Motta , Luis Rayo, Mar Reguant, Esteban Rossi-Hansberg, Sandro Shelegia, Alexandra Spitz-Oener, Claudia Steinwender, Nancy Stokey , John Sutton and participants of the UPF and University of Mannheim seminars for useful comments. All errors are ours.

1 Introduction

Recently, there has been renewed interest in apprenticeship programs across many countries.¹ Such programs can reduce the skill mismatch between firms and workers and improve labor market opportunities, especially for young and non-university educated workers (Zimmermann et al. 2013). Firms play a prominent role in these apprenticeship programs providing training and transferring knowledge to apprentices. However, there is a long-standing theoretical literature recognizing that firms are likely to provide suboptimal training (Becker 1964, Acemoglu and Pischke 1999, Garicano and Rayo 2017). In response to this classic dilemma, many governments have set up apprenticeships programs. Germany is a prime example with its dual educational system, where apprenticeships are widespread. Around 2/3 of young German workers and 1/3 of firms are involved in these programs. Both developed and developing countries, including Austria, Switzerland, France, U.K., Canada, India, Pakistan, Turkey, Colombia have active apprenticeship programs.² Despite their ubiquity, evidence on the effect of these apprenticeship programs on firms and their welfare consequences is scarce.

In this paper, we study the effects of apprenticeship programs on the labor and production decisions of firms. To do so, we use a novel administrative data set on the universe Colombian manufacturing firms with at least 10 employees, matched with rich firm-level census data between 1995 and 2009. We exploit a unique change in apprentice regulation. The change in regulation mandates firms to hire apprentices, imposing discontinuous apprentices quotas depending on firm size. To incentivize firms to hire these apprentices, the regulation also reduces the minimum wage for apprentices.

We begin by showing that the reform is successful in increasing training, roughly tripling the average number of apprentices per firm. However, the policy induces firm size distortions and it has sizeable effects on the labor and production decisions of firms. The effects are strongly heterogeneous across firms in different sectors: Firms in sectors with high skill requirements avoid training apprentices, while firms in low-skill sectors try to hire as many apprentices as possible. This heterogeneity is based on four empirical results. First, we use bunching methods in order to gauge firm size responses to the discontinuities in apprenticeship quotas. We find that firms in high-skilled sectors reduce their size in order to locate just below the regulation thresholds where they have to hire fewer apprentices, leaving missing mass above the thresholds. Firms in low-skilled sectors, on the other hand, bunch just above the regulation thresholds in order to increase the number of apprentices they can hire.

Second, we show that conditional on their post-reform size, firms in high-skilled sectors tend to hire the minimum number of apprentices required, while most firms in low-skilled sectors hire the maximum number of apprentices possible given the regulation. Third, many firms in high-skilled sectors do not even satisfy the minimum apprentice quota, but rather pay fees to the government as a “buy-out” from the regulation. In contrast, this behavior is virtually inexistent among firms in low-skilled sectors. Fourth, we find that firms in low-skilled sectors reduce the number of regular

¹We follow Wolter and Ryan (2011) definition of apprenticeships as: “programs that comprise both work-based training and formal education, in most countries at upper-secondary level, and lead to a qualification in an intermediate skill, not just to semiskilled labor”.

²See Outlook (1999), Novella and Pérez-Dávila (2017) and Fazio et al. (2016). For the current discussion on the effect of apprenticeships in England see Cavaglia et al. (2017).

workers, in particular production workers, while there is no significant effect on regular workers in high-skilled workers.

In summary, the results imply that there seems to be a net benefit to training apprentices for firms in low-skilled sectors, but a net cost of hiring apprentices in high-skilled sectors. We develop a simple model featuring heterogeneous training costs to rationalize these results, quantify the effect of the policy on welfare and analyze counterfactual policies.

Inspired by Lucas (1978), we consider an economy with heterogeneous firms characterized by their managerial ability. Firms produce a homogenous good using labor from workers and apprentices. Apprentices differ from workers in that they require training to produce. If a firm hires an apprentice it has to use some of its workers' time to train him. This training is heterogeneous across sectors. The key insight of the model is that firms that train apprentices incur in the opportunity cost of spending time teaching and not producing. Therefore, when training takes more time, the opportunity cost is larger and hence firms may prefer to avoid apprentices. As training apprentices in skilled sectors take longer than in unskilled sectors, our model provides an explanation to highlighted empirical patterns.

To rationalize our empirical observations, we model the Colombian apprentice regulation under this theoretical framework. The regulation has three main components. First, it imposes apprentice quotas: a minimum and upper bound of required apprentices depending on firm size. Second, it reduces the minimum wage for apprentices. Third, it allows firms to pay fees proportional to the number of required apprentices instead of hiring these apprentices. We show, aligned with our empirical findings, that whenever apprentice wages are sufficiently low, unskilled sectors - that require short training time - hire as many apprentices as possible. To do so, some of these firms increase their size to access more apprentices, inducing bunching at the thresholds and a missing mass of firms to the left of this regulatory cutoffs. These firms will also choose the maximum number of apprentices possible at the upper bound of the regulation and none of choose to pay the fee. The effects are opposite for firms in skilled sectors, where training is more time-consuming. We show firms in these skilled sectors try to have the least number of apprentices as possible. Those firms who have apprentices, choose the lower bound of the regulation, and to avoid having additional apprentices some of these firms decide to reduce their size and bunch below the regulatory thresholds. If the fee is low enough, they even prefer to pay the fee instead of having a suboptimal amount workers and apprentices. Our theory predicts that under this policy, firms in the unskilled sector increased their production and profits relative to firms in skilled sectors, as suggested by our data.

Our model highlights that although this type of apprenticeship programs increase the number of trained apprentices, it has heterogeneous effects across sectors. It is likely it will disproportionately benefit unskilled sector firms and incentive them to hire as many apprentices as possible. This implies most apprentices will end up being trained by unskill sector firms, a possible unintended consequence of the policy. If a policymaker also wants to increase the training of apprentices in the skilled sectors, it has to consider the different training costs across sectors. This probably isn't the only unintended consequence of the policy. We extend the model and allow firms to choose not only the number of apprentices but also the training time dedicated to each apprentice. Firms now choose between hiring more apprentices or giving more training to each of the apprentices they hire. We show that the equilibrium training time increases if apprentice wages are higher.

Intuitively if apprentice wages increase it is better to give more training to each apprentice instead of hiring more. This reveals a different mechanism through which the apprentice policy affects the total amount of training: lowering the minimum wage for apprentices increases the demand for apprentices, but reduces the amount of training each of them gets. We also show that when the wage of workers increases, the opportunity cost of teaching apprentices also rises, making optimal for firms to provide less training. This reinforces the idea that there are heterogeneous effects of the policy across sectors. If some sectors have higher wages, for instance, because their workers perform complex more productive tasks, then whenever firms hire apprentices it would give them less training.

Our model allows us to quantify these margins and the effects of the policy on welfare outcomes. Instead of taking a side on the specific functional form we exploit the idea that most welfare functions would try to maximize training and output, and reduce unemployment. We focus our welfare analysis on quantifying the change of these three variables and how they vary across sectors. At this stage, we are in the process of calibrating the model using the observed changes of the firm size distribution (bunching and missing mass), following a methodology similar to Garicano, Lelarge, and Van Reenen (2016). Using our rich micro data set we are able to target additional moments, including the number of apprentices, wages of workers by sector, production, and profits of firms. We will use the calibrated model to quantify the change in aggregate output, unemployment and on training and can compare them to counterfactual policies, such as imposing sector-specific regulations.

Related Literature: There is a rich theoretical literature examining training and human capital accumulation of workers in firms. In his seminal work on human capital Becker (1964) shows that contrary to the common wisdom of the time, firms could still provide general human capital training to workers given these workers are willing to pay for it. Similarly in our framework, if there are no minimum wage restrictions, wages would adjust to compensate for the training apprentices get. In this case, firms would provide the socially optimal training. Of course imposing a binding minimum wage prevents this from happening. Using our data we are able to quantify the welfare consequences of relaxing the minimum wage restriction, as it was the case for the change in apprentice policy.

Firms might still provide socially inefficient training. Acemoglu and Pischke (1998, 1999) show under imperfect information and imperfect competition firms have incentives to provide some general training, however training might still be suboptimal. More recently Garicano and Rayo (2017) show in a dynamic framework, that experts have the incentive to transfer knowledge inefficiently slow to apprentices. Moreover Fudenberg and Rayo (2017) argue apprentices could end up doing menial tasks and exerting inefficiently high effort on those tasks. This literature shows firms training might be insufficient so policies increasing training could be desirable. Our paper highlights that well intended apprenticeship policies that increase training, could have sizeable costs in terms of the labor decision of firms, affecting aggregate output and the composition of training. Using our rich micro data we can quantify these effects over welfare-relevant variables.

Some part of the literature has focused on explaining quality differences in apprenticeship training. For instance, Soskice (1994) shows that larger firms pay more and therefore can attract better apprentices. This discussion is relevant in our context because if differences in quality training exist between firms of skilled and unskilled sectors, these differences can explain some of our results. For instance, the quality of training provided by skilled sectors is lower and therefore students enrolled

in apprentice systems choose to only apply to unskilled sectors. To exclude this possibility, we build from Göggel and Zwick (2012) who show that differences in apprenticeship quality are usually small between firms that differ by size, sector and location.

Fersterer et al. (2008) study the effect of labour unions on the training in the German economy. Unionized firms provide better wage contracts to their employees by providing higher wage floors, incentive their apprentices to perform better with the promise to become one of these employees. We provide empirical evidence that alternative channels to wage compression can also increase training in the economy.

Most of the empirical literature on apprentices such as Krueger and Pischke (1995), Fersterer and Winter-Ebmer (2003) and Fersterer et al. (2008) focus on estimating the returns of apprenticeship programs concluding that the returns are similar to those of other types of schooling. We complement their results by studying the effects of apprentice programs on firms and not on workers.

Two closely related papers deserve more attention. First, Ospino (2018) is the only other paper that evaluates the change in the Colombian apprentice regulation. He uses census data to study the effect of this policy over the performance of firms. Using a regression discontinuity approach Ospino shows positive effects on output per worker, total factor productivity and share of exported sales, and negative effects on the wage bill of directly hired workers. We depart from Ospino's paper in three important ways. First, all the empirical facts we document are new since our paper uses administrative data collected directly by SENA to implement the apprentice policy. Moreover the census data on workers might be problematic for empirical exercises as it is not intended to capture the exact definition of the number of workers applicable for the policy. Second, our analysis shows that to understand the effects of the policy on firm's performance it is important to consider the heterogeneity in the responses of firms in skilled and unskilled sectors. Third, given we develop a structural model we are able to quantify the welfare consequences and analyze counterfactual policies.

Finally our paper is related to the literature that quantifies the effect of firm size distortion on the economy. Garicano, Lelarge, and Van Reenen (2016) study how firm size policies in France affect labor allocation and the productivity distribution. They show how employment protection laws imposed to firms with at least 50 employees, implied sizeable welfare losses as high as 5% of GDP. Similar to their paper we analyze a size dependant regulation and find substantial labor and production responses of firms. However, the apprentice policy we consider, highlights a different mechanism to that of previous studies. Apprentices in contrast to taxes and other labor regulations, are part of the firm's productive inputs. We show how this difference is important in the sense not all firms change the labor decision to avoid the policy. Firms in unskilled sectors, where training costs are lower, benefit directly from the policy increasing their size and hiring more apprentices. This unintended consequences on the composition of trained apprentices enriches the welfare considerations for the case of this apprenticeship programs. We also follow Garicano, Lelarge, and Van Reenen (2016) in our calibration methodology, extending it to include other moments from a firm-level micro data.

The rest of this document is organized as follows. In the next section, we give a brief description of our data and introduce relevant institutional context. Then we present the empirical facts of the behavioral response of firms after the apprentice regulation. We also present a preliminary empirical strategy and argue there are sizeable effects of the regulation labor and production decision of firm.

Next, we describe a simple model that rationalizes our empirical findings. Finally, we describe a preliminary methodology to calibrate the model and conclude highlighting missing pieces and future work.

2 Data and Institutional Context

We use a novel administrative data set collected by the Colombian National Apprenticeship Service (SENA) for the Colombia Ministry of Labor. It covers the universe of manufacturing firms with at least 10 workers from 1995 to 2009. There is a total of 108,385 firm-year observations, averaging approximately 7000 firms per year. For each firm-year observation we have information on the number of workers and apprentices and indicators for fees and fines paid by the firms. We match this data set to survey data from the Colombian manufacturing census (EAM). This is rich firm-level data collected by the National Department of Statistics (DANE). It includes additional information on workers divided in three production/skill layers³, detailed production inputs, wages, sales, output, and profits.

Apprentice Regulation

In December 2002 there was a major labor reform in Colombia that included a redefinition of the apprenticeship contract.⁴ The purpose of this redefinition was to increase the apprenticeships. The reform came into effect on June 2003. It defined two phases of apprenticeships, the *teaching* phase where apprentices are trained full time in an occupation and the *productive* phase where they work at the firm. Apprenticeships are of at least 6 months and have a maximum duration of 2 years. Apprentices consist of the university, technical, vocational and secondary education students or individuals with ‘semi-qualified’ formation. They have to be 14 years or older. SENA provides most of the teaching phase training to apprentices, although other tertiary education institutions can also provide this service, albeit upon SENA’s authorization. SENA is also responsible for gathering all the information on firms and apprentices and making sure both comply with the regulation. It provides a centralized matching system where firms hire apprentices. All firms (except in the construction sector) with more than 10 workers have to report the total number of permanent workers to SENA. They verify the information comparing it to other administrative data and inform firms of the applicable apprentice quota.⁵

The apprentice regulation has three important components. First, it establishes *apprentices quotas* conditional on the number of permanent workers of firms. These quotas set a minimum number

³The census data divide workers into three layers. Professional workers which are workers with tertiary education working on production. Production workers are the ones with less than tertiary education working on production and administrative workers which do not work directly in the production processes.

⁴This labor reform has other major changes in the labor regulation and was motivated in part to reduce the high levels of informality in the Colombian labor markets. None of the other policies depended on firm size. For more information on the labor reform see Law 789 of 2002 available at http://normograma.sena.edu.co/normograma/docs/ley_0789_2002.htm#12.

⁵From 2010 onwards, the information on the number of workers is compared to social security worker information (PILA) filed by the employer. For the period we study 1995-2009, SENA verified the information using census data and in-situ visits.

of apprentices that firms should have: at least one apprentice for every 20 workers, with one more for each fraction of 10 workers (divide number of workers by 20 and approximate to the nearest integer). For firms with workers between 15 and 20, the regulation stipulates they need to have at least one apprentice. The quota also sets an upper bound for apprentices. The maximum number of apprentices permitted was at most twice the minimum bound.⁶ Firms with less than 15 workers could have at most one apprentice. Second, to incentive firms to hire apprentices the regulation stipulated that non-university apprentices could be paid at least 50% of minimum wage during the teaching phase and 75% of the minimum wage during the productive phase.⁷ Third, firms could pay a fee instead of hiring the minimum required apprentices. If no apprentices are hired the fee is computed as 5% of the total number of workers multiplied to the minimum wage. A partial monetization is also possible, and it is proportional to the number of apprentices required but not hired. If the apprentice regulation is not satisfied by the firm, SENA imposes a fine equivalent to two minimum wages per apprentice not hired, plus moratory interest.

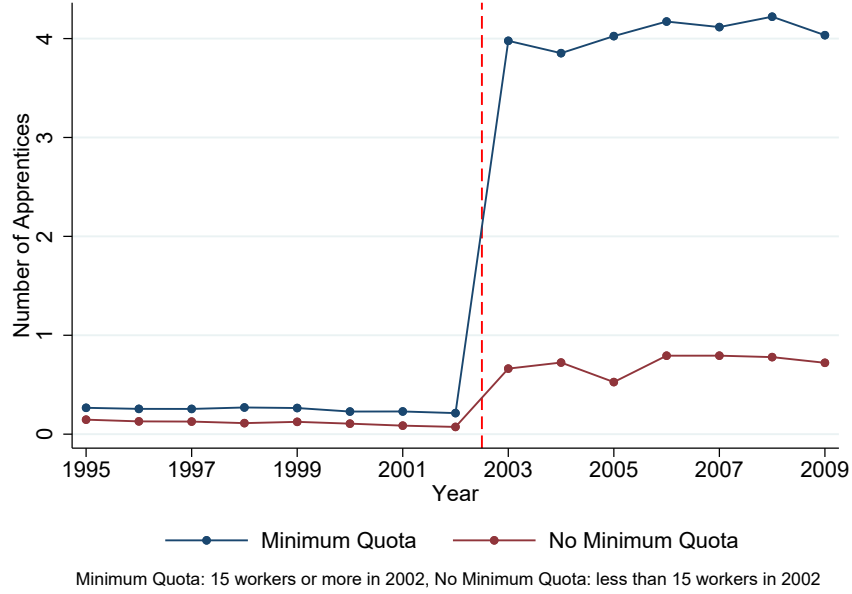
3 Number of Apprentices and Firm Size Distribution

The objective of the apprentice regulation was to increase the number of apprentices trained by firms. Figure 1 shows the policy was successful in this dimension, substantially increasing the number of trained apprentices per firm rose from 0.18 before 2003, to around 2.5 after 2003. In the figure, we split the firms into two groups: firms less than 15 permanent workers in 2002 and firms with 15 employees or more in 2002. The idea is to capture the effect of the policy on firms that were affected by minimum bound of the regulation. Although before the policy this group demanded more apprentices, after the policy the difference between the two groups amplified.

⁶This upper bound was modified in 2009. Firms between 1-14 workers could have at most 50% of total workers, 15-50 workers up to 40% of total workers, 51 to 200 workers until 30% of total workers and more than 200 workers until 20% workers (see Decree 1779 of 2009 available in http://normograma.sena.edu.co/normograma/docs/decreto_1779_2009.htm).

⁷University apprentices are paid a full minimum wage in both phases. The regulation also specified that if the unemployment rate fell below 10%, then apprentices have to be paid a full minimum wage. This only happened after 2013.

Figure 1: Average Number of Apprentices

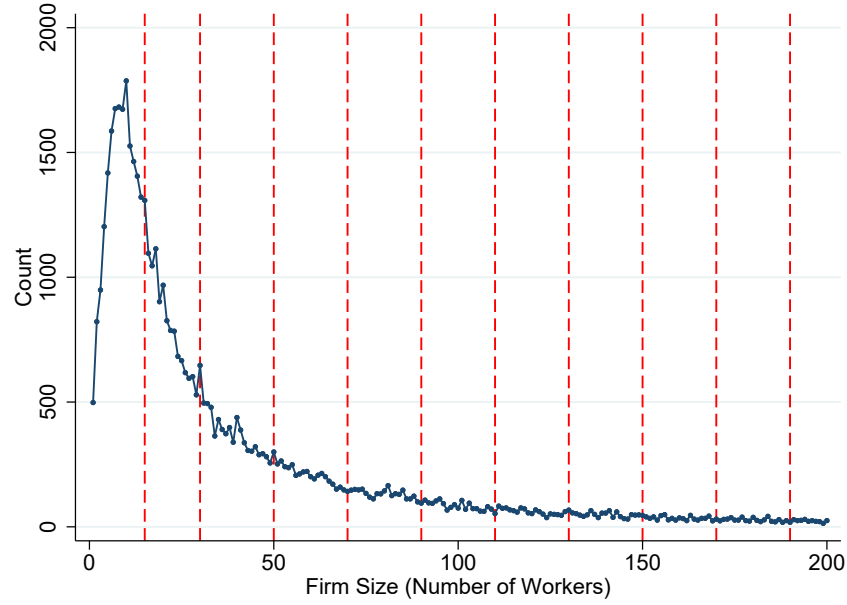


Source: Data Prepared for Ministry of Labor by the National Apprenticeship Service (SENA).

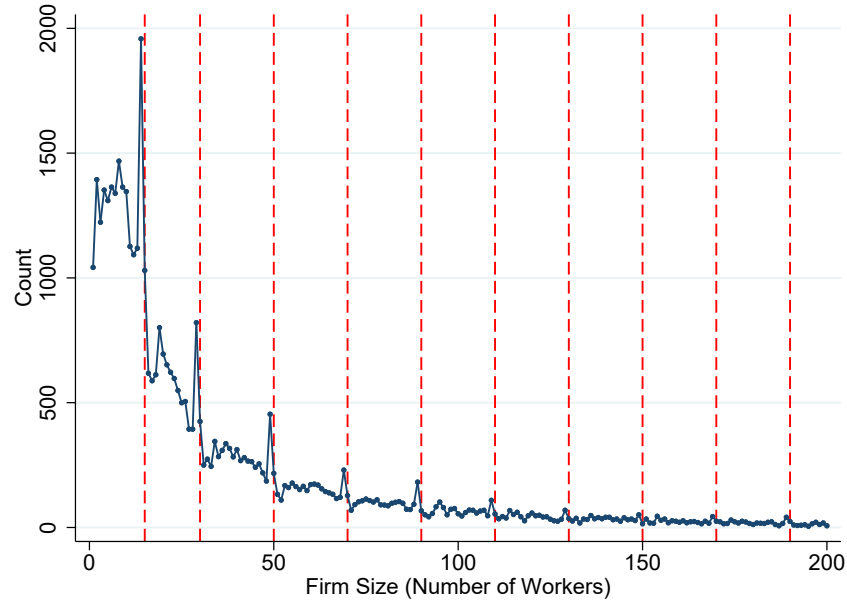
However, the regulation also induced firm size distortions by setting apprentice quotas conditional on the number of permanent workers. Figure 2 shows these distortions were sizeable and affected the reported firm size distribution after 2002.⁸ Policy thresholds are represented as the vertical dotted lines in both plots. Panel 2a shows the size distribution of permanent workers before the policy was relatively smooth. In contrast, Panel 2b shows a rugged firm size distribution, with a mass of firms ‘bunching’ before the regulatory thresholds. The dents of ‘missing mass’ around these thresholds suggest some firms change their reported size as a response to the apprentice regulation. The fact that there is missing mass both below and above the threshold indicates heterogeneous responses of firms. The missing mass above the threshold together with the bunching right below the threshold, points to some firms avoiding apprentices by reducing the number of workers they hire. In contrast, the missing mass below the thresholds, suggests some firms increased their workers to get more apprentices.

⁸In Appendix A we present the distribution every two years for the period before and after the policy. These year by year distributions have the same behavior as the pooled version.

Figure 2: Firm Size Distribution Before and After Regulation



(a) Pre Reform (1995-2002)



(b) Post Reform (2003-2009)

Source: Data Prepared for Ministry of Labor by the National Apprenticeship Service (SENA).

To understand these heterogeneous effects we first divide the manufacturing firms into nine two-

digit sectors of the Colombian industrial classification.⁹ We rank industries by the fraction of professional workers (with tertiary education) over total workers. We interpret this fraction as a proxy for the skill requirements in each sector. Sectors with more complex technologies that solve difficult tasks, probably have a larger fraction of professional workers. Whereas, sectors with routine manual tasks will likely have a smaller fraction of these workers. We divide these sectors in two: the *unskilled sectors*, where the fraction of professional workers is below the total mean of professionals over total workers and the *skilled sectors*, where fraction of professional workers above this mean. Using this definition, four sectors (Wood Products, Textiles, Food and Beverage, Mineral Non-Metallic Products) are classified as unskilled and five sectors (Paper and Editorial, Other Manufacturing, Machinery and Equipment, Metallic Products, Chemical Products) as skilled.¹⁰ This split is relatively balanced. Unskilled industries represent 48% of the firms in the whole sample (44% in the post-regulation period). In the next section, we show this way of dividing the data captures large behavioral differences across firms in response to the apprentice regulation.

4 Empirical Analysis

In this section, we document the different firm responses to the apprentice regulation. We show firms in skilled sectors avoid training apprentices, while firms in unskilled sectors want to train as many apprentices as possible. Some of the firms in the skilled sectors reduce their size below the regulatory thresholds. Whenever these firms decide to have apprentices they usually choose the lower bound of the regulation. There are some firms that even prefer to pay the fee instead of having apprentices, making the average number of apprentices they train lay below the minimum bound of the regulation. In contrast, firms in unskilled sectors seek apprentices. They increase their size and bunch at regulation thresholds to access more apprentices. They train apprentices close to the maximum bound and virtually never pay the fee.

Fact 1 : Firms in skilled sectors bunch below the thresholds, firms in unskilled sectors bunch at the thresholds. Figure 3 shows the firm size distribution for firms in skilled sectors (Panel 3a) and in unskilled sector (Panel 3b). From this figure, it is clear that firms in the skilled sectors bunch below the threshold, while firms in the unskilled sector bunch at the threshold. More formally we follow Saez (2010) to compute the bunching estimates. For each firm size distribution, we fit a 7th-degree polynomial excluding the bunching region to construct a smooth counterfactual distribution (solid line in plots). We then compute the *excess mass* $b = B/h_o(\hat{n})$ as the count of firms B at the bunching point $\hat{n}$ over the counterfactual count $h_o(\hat{n})$ at that point. This excess mass indicates the marginal bunching firm in the skilled sectors had around 2 more workers, $b \approx \Delta n \approx 2$. These estimates are relatively stable across thresholds (Panel 3a). In contrast, for unskilled sectors, the marginal bunching firm has between 0.56 and 1.79 fewer workers depending on the threshold (Panel 3b). Similarly, we can compute the *missing mass* as the ‘hole’ in the observed distribution

⁹We use DANE’s industrial classification, CIU 3 A.C, which is adapted from the International Standard Industrial Classification (ISIC) the industry classification system of the United Nations. This classification is closely related to the US Standard Industry Classification.

¹⁰In Appendix B we plot the fraction of professional workers for each of the ranked sectors and show this way of splitting the data captures a natural break-point between sectors.

relative to the counterfactual, $m = M/h_o(\hat{n})$. We compute this missing mass for a size bin of three workers. Note the missing mass has similar value to the bunching estimate suggesting those where the firms that bunched. Both the excess mass and the missing mass are statistically significant. Standard errors are presented in the plots in parenthesis and are computed by bootstrapping the firm size distribution 500 times.

Fact 2: Firms in skilled sectors choose the lower bound of the regulation, firms in unskilled sectors choose the upper bound. Figure 4 shows the lower and upper bounds of the apprentices regulation, as well as the average number of apprentices by firm size for both skilled and unskilled sectors. Firms in skilled sectors have an average number of apprentices below the regulation lower bound (Panel 4a). This implies that some of these firms are either paying the fee or not complying with the regulation. Fact 3 below shows a large fraction of firms in skilled sectors prefers to pay the fee instead of hiring the required apprentices. Of the firms in the skilled sector that hire apprentices, 86% choose the lower bound. In contrast, firms in the unskill sectors often choose the upper bound. Panel 4b shows the average number of apprentices is close to this upper bound. This means most of the apprentices are trained by firms in the unskilled sectors. Even though 56% of the observations after the reform are from firms in the skilled sector, 77% of the apprentices are trained in unskilled sector firms.

Fact 3: The majority of firms in the skilled sectors pay the fee, almost none in the unskilled sectors. Figure 5 shows the policy compliance of firms in the skilled sectors and in the unskilled sectors. In the skilled sectors around 60% of firms comply by paying the fee instead of having the required apprentices (Panel 5a). In contrast, unskilled sector firms rarely pay the fee (Panel 5b). Not a lot of firms end up paying fines, although it is relatively more frequent for skilled sector firms.

Together these three facts show firms has strong heterogeneous responses to the policy. These responses were opposite for firms in skilled and unskilled sectors. Firms in the skilled sectors tried to avoid training apprentices, while firms in unskilled sectors wanted more apprentices. Though the policy was successful in increasing the number of trained apprentices, it did so at the expense of inducing labor distortions. Moreover, most of the apprentices were trained by unskilled sector firms, something that isn't necessarily desirable or intended by the policy.

To study the consequences of these distortions on other firms outcomes, we run empirical specifications looking at the effects of the policy on firm output and profits.

Fact 4: Effect on labor composition, productivity and profits (in progress)

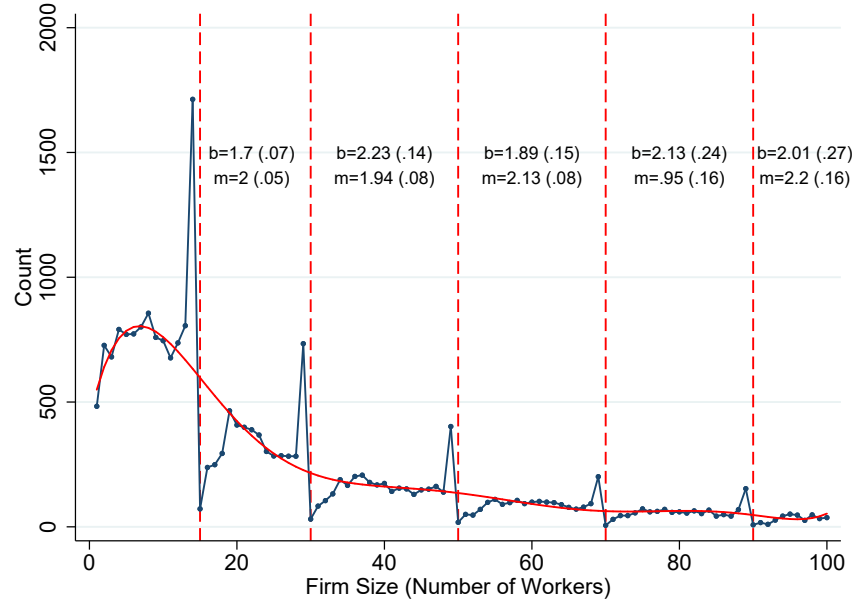
i. Dif-in-dif strategy

To study the effects of the policy on further outcomes, we employ the following difference-in-difference specification:

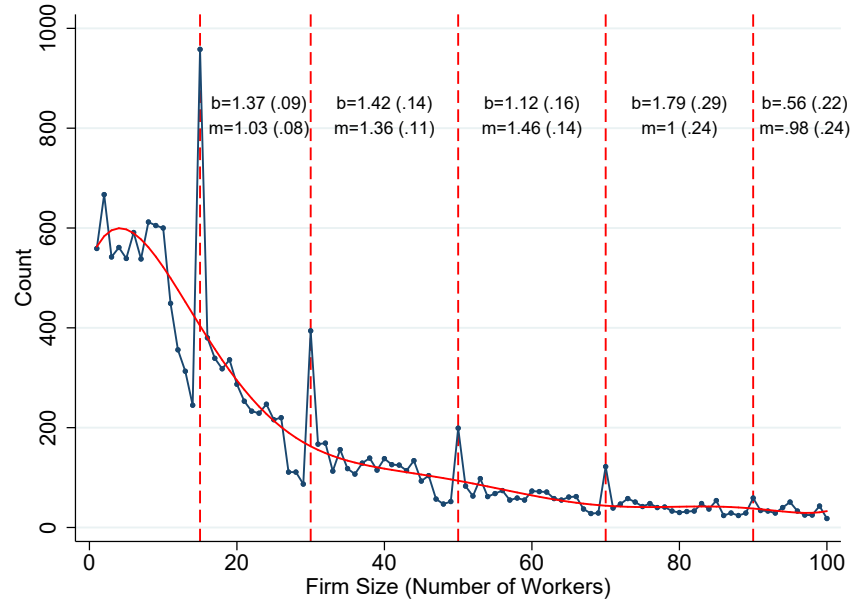
$$Y_{it} = \beta_0 + \beta_1 treat_i + \beta_2 post_t + \beta_3 treat_i post_t + X'_{it} \gamma + \epsilon_{it} \quad (1)$$

where $treat_i$ is an indicator for the treatment group, $post_t$ is an indicator for the post-reform years from 2003 onwards, X_{it} is a vector of controls and ϵ_{it} is an error term.

Figure 3: Fact 1 - Bunching of Firms in Skilled and Unskilled Sectors



(a) Skilled Sectors

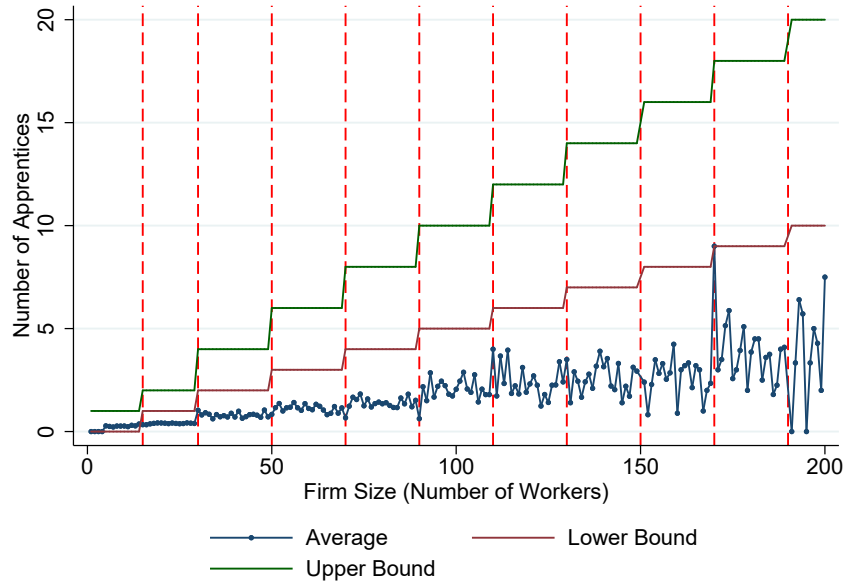


(b) Unskilled Sectors

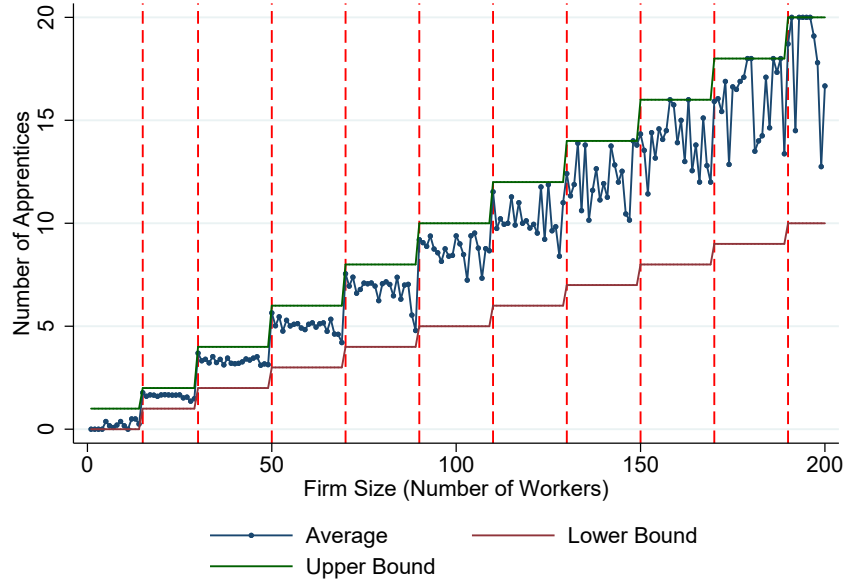
Source: Data Prepared for Ministry of Labor by the National Apprenticeship Service (SENA).

Treatment is defined based on the firm's size in 2002, the last pre-reform year. Firms with 15 or more workers in 2002 are subject to the minimum apprentice quota and serve as the treatment

Figure 4: Fact 2- Apprentices and Number of Workers



(a) Skilled Sectors



(b) Unskilled Sectors

Source: Data Prepared for Ministry of Labor by the National Apprenticeship Service (SENA).

group, while firms with less than 15 workers serve as the control group. The key advantage of assigning treatment based on pre-reform size is that the pre-reform firm size distribution is not

Figure 5: Fact 3- Policy Compliance



(a) Skilled Sectors



(b) Unskilled Sectors

Source: Data Prepared for Ministry of Labor by the National Apprenticeship Service (SENA).

subject to manipulation, i.e. assignment to treatment is exogenous. In addition, the treatment group is limited to firms with 15 to 29 workers in order to keep the two groups as similar as

possible.

Panels 6a and 6b show the “first stage” of this specification. Firms in the treatment group experience a significant increase in the number of apprentices by around 0.6 after the reform, while pre-trends are flat. At the same time, firms treated by the minimum quota are about 20% more likely to pay fees to the government after the regulation. Panel 6c of figure 6 shows the effect of on one outcome, namely the total number of workers in the firm (excluding apprentices). Treated firms seem to reduce the number of regular of workers in response to hiring more apprentices. Panels 6d and 6e show that this effect is driven by a reduction in the number of production workers, while the number of professionals increases slightly if anything. We are currently working on adding further outcomes to this analysis, in particular regarding the effects on productivity and profits.

ii. Triple-dif strategy

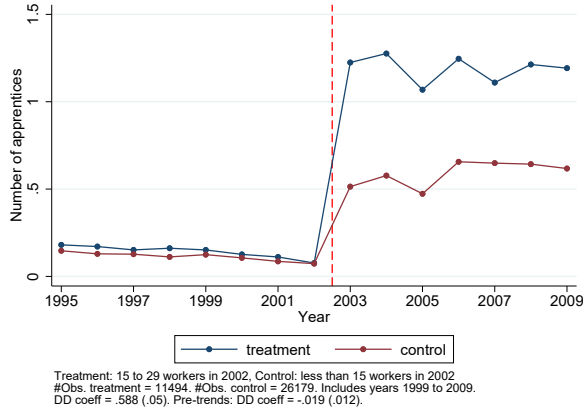
The dif-in-dif strategy may be useful in gauging the overall effect of the policy, but it has two drawbacks. First, it relies on comparing firms of different size who may not necessarily exhibit parallel trends in key outcomes. For instance, smaller firms may be growing faster than larger firms. Second, the basic dif-in-dif strategy does not speak to the differential effects of the regulation on firms in skilled vs. unskilled sectors. To overcome these two weaknesses, we run the following triple-dif specification:

$$Y_{it} = \beta_0 + \beta_1 treat_i + \beta_2 post_t + \beta_3 skill_i + \beta_4 treat_i post_t + \beta_5 skill_i post_t + \beta_6 skill_i treat_i + \beta_7 skill_i treat_i post_t + X'_{it}\gamma + \epsilon_{it} \quad (2)$$

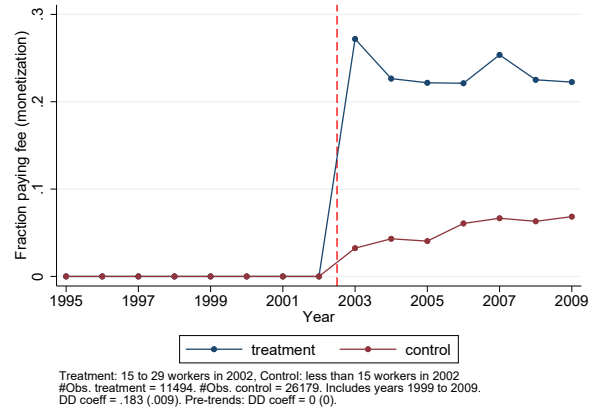
where the additional variable $skill_i$ is an indicator for firm i being in a skilled sector. The coefficient of interest is now β_7 , capturing the differential effect of the regulation on treated firms in skilled sectors. Note that this specification allows both for differential trends across firms of different size (treatment vs. control) and for differential trends across sectors (skilled vs. unskilled).

Panels 7a and 7b show the “first stage” of this specification. Treated firms in all sectors take in more apprentices as a result of the reform, but the magnitude of the effect is much larger in unskilled sectors. Conversely, there is a large fraction of treated firms paying fees to the government in skilled sectors, whereas even treated firms in unskilled sectors virtually never pay fees. Panel 7c focuses on the effect on the number of regular workers in the firm. There is no significant effect in skilled sectors, but firms in unskilled sectors reduce the number of regular workers. Panels 7d and 7e show that the effect is almost entirely driven by a reduction in the number of production workers. We are currently working on adding further outcomes to this analysis, in particular regarding the effects on productivity and profits.

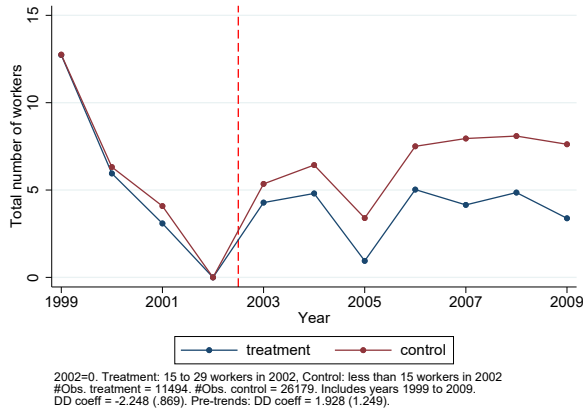
Figure 6: Fact 4- Difference-in-difference estimation



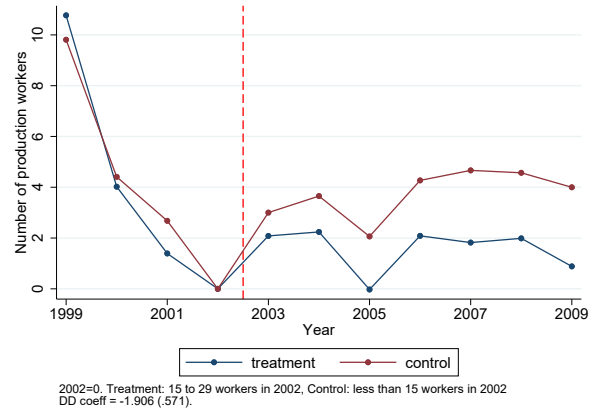
(a) Number of apprentices



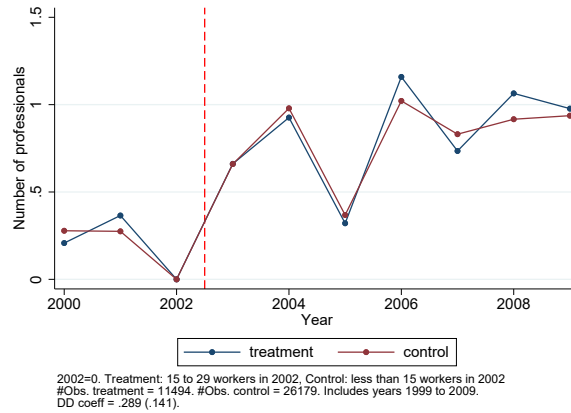
(b) Fraction paying fee



(c) Total number of workers

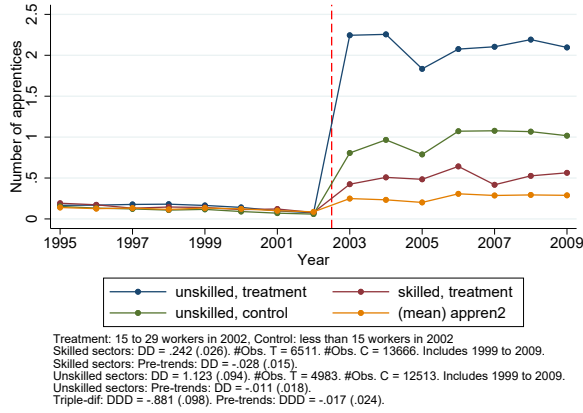


(d) Production workers

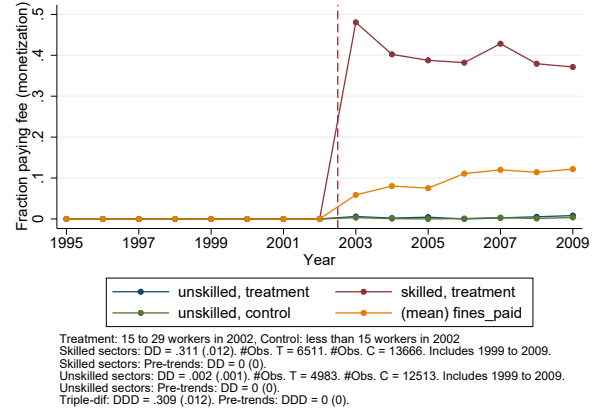


(e) Professionals

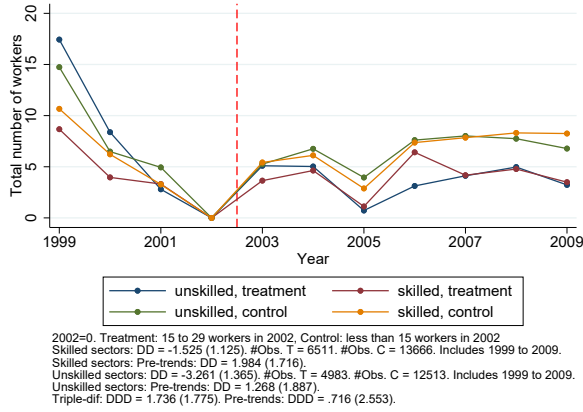
Figure 7: Fact 4- Triple-difference estimation



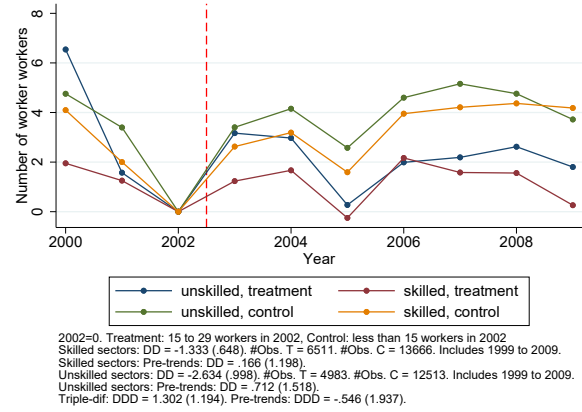
(a) Number of apprentices



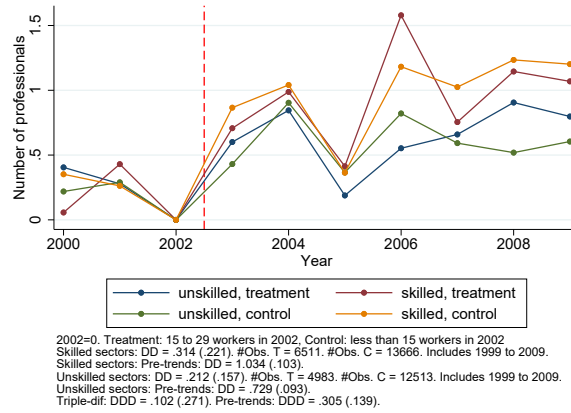
(b) Fraction paying fee



(c) Total number of workers



(d) Production workers



(e) Professionals

5 Model

In this section, we present the baseline model that captures our empirical predictions and will allow us to quantify the effect of the policy on welfare. Later we will use the model to study counterfactual policies and compare the effect of welfare-relevant variables. This first model is static. For exposition purposes, we describe production in one sector, the extension to multiple sectors is immediate and it is covered below.

5.1 Without Regulation

Suppose an economy with heterogeneous firms characterized by a managerial ability $z \sim \Phi(z)$, with a continuous density $\phi(z)$ (similar to Lucas 1978). This managerial ability is an idiosyncratic characteristic of the firm that can also be interpreted as technological differences or other factors that affect a firm's productivity. Firms in each sector produce a homogenous final good $y(z)$ using labor.¹¹ Labor is supplied either by trained workers l or by *apprentices* l_a . Suppose both types of individuals have a unit of time which they supply inelastically. Production combines firm's managerial ability z , with the total labor supplied by both type of workers, $f(l, l_a; z)$. More labor from workers or apprentices, or higher managerial ability produce more output, $f_l, f_{l_a}, f_z > 0$.¹²

Apprentices are trained using workers' time. Let $t_a \in [0, 1]$ denote the units of time a worker spends training apprentices. Once trained, an apprentice supplies $g(t_a)$ units of labor. The function $g : [0, 1] \rightarrow [0, 1]$ denotes the training technology and can be interpreted as how difficult it is to teach/learn a particular task. First we assume t_a is exogenous and set $g(t_a) := 1$. Later we endogenize the firm's training decision. In this first case, larger exogenous t_a means it takes more time to fully train an apprentice. Note that depending on the way the production function combines both types of labor, the labor inputs from workers and apprentices can be complements or substitutes in production. Apprentices could be trained in tasks that substitute workers' labor or tasks that complement it.

Firms choose the number of workers and apprentices to maximize their profits. If a firm hires n workers and train n_a apprentices, the total labor supplied by the workers is $l := n - t_a n_a$ and by the apprentices $l_a := g(t_a) n_a$. Firms interact in a competitive market, taking the wage of workers w and of apprentices w_a as given. We normalize the price of the final good to 1.

A firm with managerial ability z solves,

$$\max_{n, n_a \geq 0} f(n - t_a n_a, g(t_a) n_a; z) - wn - w_a n_a \quad s.t \quad t_a n_a \leq n. \quad (3)$$

The constraint implies the number of workers has to be large enough to train the total number of apprentices.

¹¹The model could be readily extended to have capital or other production inputs. Here we suppose labor is the only input to simplify the model and emphasize the role of apprentices.

¹²We use the notation f_x to denote $\frac{\partial f}{\partial x}$.

We can substitute l and l_a to write an equivalent (perhaps more familiar) optimization problem,

$$\max_{l, l_a \geq 0} f(l, l_a; z) - wl - \frac{w_a + t_a w}{g(t_a)} l_a.$$

We make some simplifying assumptions on f to further characterize the optimal number of workers and apprentices, guaranteeing their existence and uniqueness.

Assumption 1. (*Production Function*) Suppose $f : \mathbb{R}_+^3 \rightarrow \mathbb{R}_+$ is twice continuously differentiable and

- (i) homogenous of degree $\gamma \in (0, 1)$ in (l, l_a) .
- (ii) the Inada conditions hold, $\lim_{x \rightarrow 0} f_x = \infty$ and $\lim_{x \rightarrow \infty} f_x = 0$ for $x \in \{l, l_a\}$.
- (iii) has non-negative cross derivatives with respect to z , i.e. $f_{l,z} \geq 0$ and $f_{l_a,z} \geq 0$.

Condition (i) and (ii) imply the existence of a unique interior solution. Condition (iii) that the optimal number of workers $n^*(z)$ and apprentices $n_a^*(z)$ are non-decreasing in z . This means firms with higher managerial ability are larger. We formalize these claims in Lemma 1.

Lemma 1. *Assumptions 1 imply there are unique labor demands $n^*(z), n_a^*(z) > 0$, with $t_a n_a^* < n^*$ solving the firm z 's optimization problem (3). Moreover, these labor demands are non-decreasing in the firm's managerial ability, $\frac{\partial n^*(z)}{\partial z} \geq 0$ and $\frac{\partial n_a^*(z)}{\partial z} \geq 0$.*

We can further characterize the solution looking at the FOCs of (3),

$$[n] : \frac{\partial f}{\partial l} = w \quad , \quad [n_a] : \frac{\partial f}{\partial l_a} g(t_a) = w_a + t_a w.$$

Intuitively the marginal cost of an apprentice is not only its wage w_a but also t_a units of time of a worker that earns wage w . The firm optimizes where the marginal product of an additional apprentice is equal to this marginal cost, $w_a + t_a w$. In equilibrium, the marginal rate of substitution between the two types of labor is equal to the ratio of marginal labor costs,

$$-\frac{f_l}{f_{l_a} g(t_a)} = -\frac{w}{w_a + t_a w}.$$

We can use this equation to think how changing wages or the required training time affect the optimal decision. As usual, an increase in the relative wages of apprentices decreases their demand. In the case where t_a is exogenous and $g(t_a) \equiv 1$, an increase in training time decreases the demand for apprentices.

Lemma 2. *Suppose Assumption 1 holds and the training time is exogenous (with $g(t_a) \equiv 1$), then $\frac{n_a}{n}$ is decreasing in w_a and t_a , and increasing in w .*

5.1.1 Competitive Equilibrium

We suppose a simple supply side to this economy. There is a fixed mass of workers L and of apprentices L_a . Workers are able to do the work of apprentices but not the other way around. This implies that in equilibrium the apprentice wages has to be smaller or equal to the one that workers get, $w \geq w_a$. We consider a minimum wage that could be binding. So together both constraints imply, $w \geq w_a \geq w_{min} \geq 0$.¹³

Let $N := \int n^*(z)\phi(z)dz$ denote the aggregate demand for workers and $N_a := \int n_a^*(z)\phi(z)dz$ the aggregate demand for apprentices. The market clearing conditions are,

$$N + U = L \quad , \quad N_a + U_a = L_a, \quad (4)$$

where $U, U_a \geq 0$ denote the ‘unemployed’ workers or apprentices.

Definition 1. *A competitive equilibrium is given by wages (w^*, w_a^*) , unemployed workers and apprentices (U^*, U_a^*) and labor demands $(n^*(z), n_a^*(z))$ for each firm z , such that,*

- (i) *firms solve the optimization problem (3),*
- (ii) *wage restrictions are satisfied, $w^* \geq w_a^* \geq w_{min}$ and*
- (iii) *labor markets clear (4) .*

Notice there is only unemployment when the wage restrictions are binding. If there are some unemployed workers $U^* > 0$, then $w^* = w_a^*$. Similarly if there are some unemployed apprentices $U_a^* > 0$ then $w_a^* = w_{min}$. The case where these restrictions are binding changes the welfare consequences of the labor policy and the model predictions. We spell out this in the sections below.

5.2 Regulation

In this section, we describe the firm-size policy within the described framework. This policy intended to increase the number of apprentices trained in two ways. First, it reduced the effective lower bound firms pay for an apprentices to $w_a \geq w_{min}^a$, where w_{min}^a is lower than the minimum wage w_{min} .¹⁴ Second, it imposed apprentice quotas based on the firm-size (number of workers). If the number of workers is $n \in [N_{j-1}, N_j)$ then the number of apprentices has to be $n_a \in [\underline{n}_a^j, \bar{n}_a^j] =: \mathcal{A}_j$, $\forall j$. Alternatively, firms can pay a fee $f_a(n, n_a)$ instead of having the required apprentices. This fee is a function of the total number of workers n and apprentices the firm chooses. It is proportional to the difference between the minimum number of required apprentices $\underline{n}_a^j$ and the apprentices hired n_a .

¹³The informal labor markets might absorb some of the workers displaced by the minimum wage regulations. In particular, in Colombia the informal labor market is large. This will be relevant for the welfare calculations when thinking about the outside options of workers when the minimum wage is binding. In our empirical analysis, we only observed the formal sector, formal firms that report formal workers.

¹⁴For the period we consider, most apprentices get paid 50% of the minimum wage in the learning phase and 75% in the productive phase.

Firm z 's optimization problem when facing this regulation is,

$$\begin{aligned}
& \max_{n, n_a \geq 0, d_{f_a} \in \{0,1\}} f(n - t_a n_a, g(t_a) n_a; z) - wn - w_a n_a - d_{f_a} f_a(n, n_a) \quad s.t. \quad t_a n_a \leq n \\
& \text{if } n \in [N_{j-1}, N_j) \text{ and } d_{f_a} = 0, \text{ then } n_a \in [\underline{n}_a^j, \bar{n}_a^j] \quad \forall j, \\
& \text{if } n \in [N_{j-1}, N_j) \text{ and } d_{f_a} = 1, \text{ then } f_a(n, n_a) = \psi_a n \frac{(n_a^j - n_a)^+}{\underline{n}_a^j}, \quad n_a \leq \bar{n}_a^j, \quad \forall j.
\end{aligned} \tag{5}$$

where $d_{f_a} \in \{0,1\}$ is the decision of paying the fee or not and ψ_a is a positive constant. When a firm decides to pay the fee, it can hire fewer apprentices than the minimum required $\underline{n}_a^j$ and pay an amount proportional to the number of workers hired multiplied to the percentage difference between the quota's lower bound and the apprentices hired, $f_a(n, n_a) = \psi_a n \frac{(n_a^j - n_a)^+}{\underline{n}_a^j}$. As in the actual regulation paying the fee doesn't allow the firm to exceed the upper bound of the apprentice quota, $n_a \leq \bar{n}_a^j$. This fee function only takes the positive difference between the lower bound of the regulation and the number of apprentices hired. We assume that there is no fee for firms that choose apprentices above the lower bound. In that case we suppose that firms choose $d_{f_a} = 0$.

Let us focus on the case where the optimal number of apprentices is a fix proportion of the labor force, $n_a^* = Bn^*$ with $B \in \mathbb{R}_+$.¹⁵ We show that if relative wages of apprentices are low enough $\frac{w_a}{w} \rightarrow 0$ and the training time is low $t_a \rightarrow 0$, firms want to get apprentices. In this case, the optimal number of apprentices is above the regulation's upper bound for firms larger than N_1 . Firms bunch at the thresholds N_j 's, with missing mass to the left and never pay the fee. In contrast when $\frac{w_a}{w} \rightarrow \infty$, the optimal number of apprentices converges to zero and lay below the regulation's lower bound for firms larger than N_1 . This implies some firms to the right of the thresholds N_j 's reduce their size and bunch just below the thresholds to avoid having extra apprentices. Additionally, if the fee is low enough (ψ_a small), then some firms prefer to pay the fee instead of having the additionally required apprentices. Proposition 1 compiles these results.

Proposition 1. *Suppose Assumptions 1 holds and firms solve the maximization problem with regulation (5). Then,*

Case 1: *there exist $(\frac{\bar{w}_a}{w}, \bar{t}_a)$ such that for $\frac{w_a}{w} \leq \frac{\bar{w}_a}{w}$ and $t_a \leq \bar{t}_a$,*

- i. the number of apprentices without regulation is $n_a^* = B_u n^*$ and lays above the regulation's upper bound, $n_a^*(n) > \bar{n}_a^j$.*
- ii. there exist cutoffs $\{z_b^j, z_r^j\}_j$ such that firms $z \in [z_b^j, z_r^j]$, increase their size to the threshold N_{j+1} , so there is missing mass to the left of the thresholds.*
- iii. firms choose the upper-bound of the regulation $n_a^r = \bar{n}_a^j$.*
- iv. firms never pay the fee.*

Case 2: *there exist $\frac{w_a}{w}$ such that for $\frac{w_a}{w} \geq \frac{w_a}{w}$,*

¹⁵Note that the homogeneity of degree γ allows for this case. For the linear relationship between apprentices and workers, we additionally need that f_l/f_{l_a} is homogenous of degree zero in z .

- i. the number of apprentices without regulation is $n_a^* = B_s n^*$ and lays below the regulation's lower bound, $n_a^*(n) < \underline{n}_a^j$.
- ii. there exist cutoffs $\{z_b^j, z_r^j\}_j$ such that firms $z \in [z_b^j, z_r^j]$, reduce their size ϵ below the threshold N_j .
- iii. firms that choose to increase apprentices, choose the lower bounds $\underline{n}_a^j$.
- iv. there exist $\bar{\psi}_a > 0$ such that for $\psi_a \leq \bar{\psi}_a$, there is an additional cutoffs z_f^j such that firms $z \in (z_r^j, z_f^j]$ choose to pay the fee.

Figure 8: Apprentices After Regulation

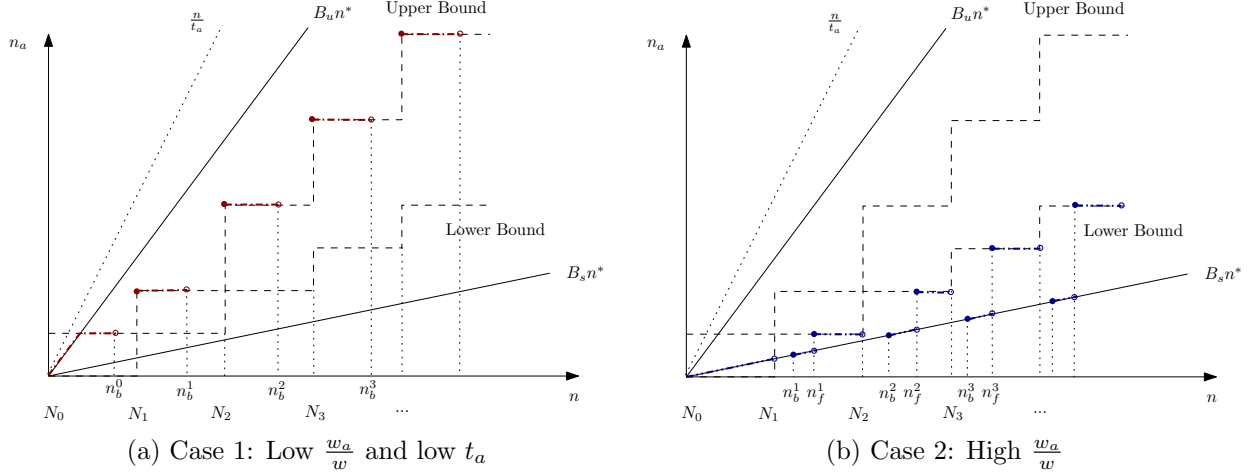


Figure 8 illustrates the two cases in Proposition 1. Panel 8a depicts Case 1, where both $\frac{w_a}{w}$ and t_a are small. Firms want to have apprentices as they are relatively cheap in terms of wages and training time. In this case, the optimal number of apprentices without regulation lays above the regulation's upper bound. This implies that some firms increase their size and bunch at the thresholds, leaving a missing mass of firms to the left of them. All of these firms choose the upper bound of the regulation. Panel 8b depicts Case 2, where $\frac{w_a}{w}$ is large. In this case, workers are more attractive than apprentices to firms. The optimal number proportion of apprentices lays below the lower bound of the regulation. Firms bunch just below the threshold to avoid having extra apprentices from the policy. There is a missing mass of firms to the right of the thresholds. If the fee is relatively small, some firms prefer to pay the fee and choose the optimal number of apprentices and workers. Firms that comply the policy with the apprentice quota choose the lower bound of the regulation.

Figure 9: Firm Size Distribution

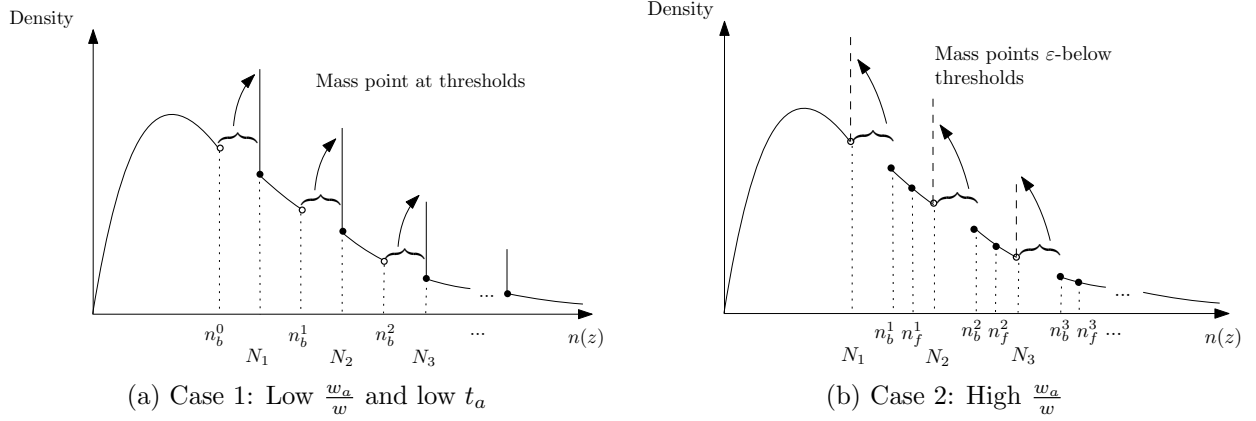


Figure 9 shows the implications of the policy on the firm size distribution for the two cases in Proposition 1. Panel 9a depicts the case where firms increase their size to get more apprentices. They bunch at each of the thresholds $\{N_j\}_j$ and there is a missing mass of firms between $[n_b^j, N_{j+1})$. Panel 9b illustrates Case 2, where firms either reduce their size or pay the fee to avoid having more apprentices. Firms bunch below each threshold, leaving a missing mass of firms of size between $[N_j, n_b^j)$. If the fee is low enough, firms of size $[n_b^j, n_f^j]$ prefer to pay the fee instead of having the required apprentices.

5.3 Effects of the Regulation

Proposition 1 gives us a framework to understand our first three empirical findings. The relaxation of the apprentices minimum wage, allowed firms in unskilled sectors to find it profitable to hire as many apprentices as possible. This implies some firms in the unskilled sector bunch at the thresholds to get more apprentices, choose the upper bound of the regulation and never pay the fee. On the other hand, even though apprentice wages went down, it wasn't enough for skilled sector firms to demand more apprentices than the lower bound of the regulation. These firms would, therefore, avoid having additional apprentices. Some of them choose to decrease their size and bunch just below the thresholds. Others prefer to pay the fee instead of having the required apprentices.

However, this proposition doesn't tell us about the effect of the regulation on firms' output and profits. To understand these effects is useful to distinguish the three components of the regulation: i) the change in the apprentice minimum wage w_{min}^a , ii) the apprentice quotas based on firm size and iii) the possibility of paying a fee. Each of these has a different effect on output and profits, implications we can later test and use empirically.

Apprentice Wages

If before the regulation the apprentice minimum wage is binding, $w_a^* < w_{min}^a$, then a decrease in w_{min}^a increases output and profits. This follows immediately from the firm's maximization problem (3). Let $\pi^*(z)$ denote the firm z 's profits and $y^*(z)$ its output from the solution of the optimization problem without regulation. Firms profits $\pi^*(z)$ and output $y^*(z)$, are decreasing in wages w_a and w . This implies that a decrease in the binding minimum wage w_{min}^a increases profits and output for all firms. Clearly, this is a Pareto improvement. Allowing firms to pay lower wages to apprentices improves efficiency.

Apprentice Quotas

Introducing apprentice quotas always decrease total profits and output. If firms want more apprentices than the quota, then the binding regulation will limit their output and lower their profits. If firms want fewer apprentices than the minimum bound required by the regulation, some would hire fewer workers to avoid the regulation reducing output. Firms' profits decrease both for these firms and the ones that decide to hire the apprentices.

Apprentice Fee

The regulation also allows firms to pay a fee instead of having the required apprentices. Firms that decide to pay the fee have the same output as in the case without regulation, but lower profits. This is the case since once they pay the fee they optimally choose the same number of workers and apprentices as they would without the regulation.

Using our structural model we can quantify the net effect of each of these margins on the total change in output and profits observed for each sector. Moreover, making use of the calibrated model we plan to study the net welfare implications in the economy.

5.4 Endogenous Training

Firms might endogenously choose how much to train apprentices. Function $g(\cdot)$ represents in this case how difficult it is to train an apprentice. Suppose more training time invested in apprentices makes them more productive $g'(t_a) > 0$, but there are decreasing return to training $g''(t_a) < 0$. To solve for endogenous training, we add to firm z 's original optimization problem (3), the choice of $t_a \in [0, 1]$.

The FOC with respect to t_a implies,

$$f_l n_a = g'(t_a) f_{l_a} n_a,$$

so the marginal cost of training an apprentice is equalized to the marginal improvement in the apprentice abilities.

If $n_a^* \neq 0$ and using the FOCs of (3),

$$\frac{g'(t_a)}{g(t_a)} = \frac{w}{w_a + t_a w}. \quad (6)$$

Additionally suppose, $\lim_{t_a \rightarrow 0} g'(t_a) = \infty$ and $\lim_{t_a \rightarrow 1} g'(t_a) = 0$. The first condition states that supplying a small amount of training significantly improves apprentices productivity. The second that once workers allocate most of their unit of time to training apprentices, more training won't increase productivity too much. These two conditions together with the concavity of g , guarantee a unique interior solution $t_a^* \in (0, 1)$ that solves the firms training decision.

Note that this training decision doesn't depend on the firm's managerial ability z . This means training per apprentice won't change for firms of different size or with different number of apprentices. Optimal training is only affected by changes in the training technology $g(\cdot)$ or wages. In particular, an increase in the apprentice wage w_a increases the optimal amount of training. Intuitively, hiring more apprentices becomes more costly so firms choose to train more each of the apprentices they hire. Conversely, an increase in workers wages w decreases the amount of training each apprentice gets. As the opportunity cost of training is the worker's wage, this opportunity cost increases as the wage of workers rise. Lemma 3 summarizes these results.

Lemma 3. *Suppose $g'(t_a) > 0$, $g''(t_a) < 0$, $\lim_{t_a \rightarrow 0} g'(t_a) = \infty$ and $\lim_{t_a \rightarrow 1} g'(t_a) = 0$. Then there exist a unique $t_a^* \in (0, 1)$ that solves (6). Moreover, $\frac{\partial t_a^*}{\partial w_a} > 0$ and $\frac{\partial t_a^*}{\partial w} < 0$.*

5.5 Multiple Sectors

The same analysis applies when we extend the model to K sectors. Suppose managerial ability is sector specific z_k and has an exogenous distribution Φ_k with continuous density ϕ_k . Firms of different sectors differ in their production $f^k(\cdot)$ and training technologies $g^k(\cdot)$. For the case where the training is exogenous, we assume that the required training time t_a^k to fully train an apprentice $g^k \equiv 1$ varies across sectors. A sector whose technology requires simple menial tasks probably requires less of the workers time to train apprentices. On the contrary, a sector with highly specialized tasks requires more training.

A firm z_k in sector k solves the same problem as in (3),

$$\max_{n_k, n_{a,k} \geq 0} f^k(n_k - t_a^k n_{a,k}, g^k(t_a^k) n_{a,k}; z_k) - w n_k - w_a n_{a,k} \quad s.t. \quad t_a^k n_{a,k} \leq n_k.$$

Let $n_k^*(z_k)$ and $n_{a,k}^*(z_k)$ denote the optimal labor demand for workers and apprentices in sector k . If Assumptions 1 hold in every sector, Lemmas 1 and 2 follow.

The main difference when solving for the competitive equilibrium with multiple sectors is the market clearing conditions. We consider two possibilities, either there is a common pool of workers who can work in either sector or separate labor markets for workers in each sector. We study both of these cases below. For apprentices, we suppose there is a fixed mass of L_a apprentices that can be trained in any sector.

First, let us suppose there is a mass L of workers that can work in any sector. This implies in equilibrium both sectors will pay the same wage rate w to workers. Let $N_k := \int n_k^*(z_k)\phi_k(z_k)dz_k$ denote the aggregate demand for workers in sector k and $N_{a,k} := \int n_{a,k}(z_k)\phi_k(z_k)dz_k$ the aggregate demand for apprentices in sector k . The market clearing conditions in the common worker case are,

$$\sum N_k + U = L \quad , \quad \sum N_{a,k} + U_a = L_a.$$

In this case a competitive equilibrium is still defined as wages (w_a^*, w^*) , unemployed workers and apprentices (U^*, U_a^*) and labor demands for each sector k , that satisfy Definition 1 in each sector.

In the second case, we suppose that the labor markets in each sector are perfectly segmented and there is a fixed supply of workers L_k in each sector. In this case, there is a separate market clearing condition and equilibrium wage rate w_k^* for each sector k ,

$$N_k + U_k = L_k \quad , \quad \sum N_{a,k} + U_a = L_a.$$

In a competitive equilibrium there are now an equilibrium wage in each sector $\{w_k^*\}_k$ and unemployment of workers in each sector $\{U_k^*\}_k$.

The same analysis of the regulation applies separately to each sector. Depending on the equilibrium wages $(w_a^*, \{w_k^*\}_k)$ and the training technologies t_a^k , different cases of Proposition 1 apply.

5.6 Welfare

We study the welfare consequences of the policy by considering its effect over total output, the number of trained apprentices and the number of unemployed workers. A benevolent social planner weights these three variables through some social welfare function. Here we don't attempt to spell out this welfare function, but rather quantify what happens to each of its 'inputs' when with the apprentice regulation.¹⁶ Notice that when we consider multiple sectors there are additional composition effects that could also be important for the social planner. For instance, it might not be socially optimal to train all apprentices in unskilled sectors. Therefore we also consider the composition consequences of the regulation.

Formally a regulation $\mathcal{R}$ consists of a set thresholds N_j 's, a set of lower and upper bounds $\underline{n}_a^j, \bar{n}_a^j$, and a fee function $f_a : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$, $\mathcal{R} := \{N_j, \underline{n}_a^j, \bar{n}_a^j\}_j, f_a\}$. Let $y_k^*(z)$ denote the output a firm in sector k with managerial ability z produces when solving (5). Aggregate output is the sum of the output of all firms across every sector,

$$Y = \sum_k \int y_k(z)\phi_k(z)dz.$$

¹⁶We can think of a dynamic model that explicitly models the effects of training workers over the future labor force as well as other positive effects of training. At this point, we abstract from these considerations in particular because we would need additional data on training and its effects on workers to discipline these margins in the quantitative exercises considered below.

We simply compute the change in aggregate output as the difference between aggregate output with regulation and without regulation, $\Delta Y = Y^r - Y^*$.

Similarly, we compute the change in unemployed workers as $\Delta U = U^r - U^*$ and unemployed apprentices as $\Delta U_a = U_a^r - U_a^*$. We also calculate and report the change in trained apprentices by sector, $\Delta N_{a,k} = N_{a,k}^r - N_{a,k}^*$.

Once we calibrate the model, we are able to compute the changes in these variables and decompose the effect of each of the elements of the regulation. Moreover, we can use the calibrated model to compare the implications of counterfactual models.

6 Quantitative Exercises (In Progress)

In this quantitative section, we plan to calibrate the model and use it to quantify the effect of the policy on the welfare variables of interest as well as analyze policy counterfactuals. We are still working on the calibration and the quantitative results of the model. Here we briefly describe the model's parametrization and our calibration strategy.

For our quantitative exercises, we parameterize the production function as a CES function. In particular, suppose,

$$f(l, l_a; z) = [\eta(z)l^\rho + \eta_a(z)l_a^\rho]^{\frac{\gamma}{\rho}},$$

where $\gamma \in (0, 1)$ and $\frac{1}{1-\rho}$ is the elasticity of substitution between the two labor inputs.¹⁷ The functions $\eta(z)$ and $\eta_a(z)$ allow some flexibility on how the managerial ability affects each type of labor. To have a linear relationship between the optimal number of apprentices and the number of workers, we assume the ratio of $\eta_a(z)/\eta(z)$ doesn't depend on z .

The FOC for the CES case (with $g(t_a) = 1$) imply,

$$\frac{l_a}{l} := \frac{n_a}{n - t_a n_a} = \left(\frac{\eta_a(z)}{\eta(z)} \frac{w}{wt_a + w_a} \right)^{\frac{1}{1-\rho}} =: A \in \mathbb{R}_+.$$

So the ratio of apprentices to workers is,

$$\frac{n_a}{n} = \frac{A}{1 + At_a} =: B.$$

This ratio $\frac{n_a}{n}$ is decreasing in t_a and $\frac{w_a}{w}$ as Lemma 2 predicts.

We can explicitly solve for the labor demand functions,

$$n_a^*(z) = \left[\frac{\gamma [\eta(z)A^{-\rho} + \eta_a(z)]^{\frac{\gamma-\rho}{\rho}} \eta(z)A^{1-\rho}}{w} \right]^{\frac{1}{1-\gamma}}, \quad n^*(z) = \frac{n_a^*(z)}{B}.$$

¹⁷If $\rho = 1$, then both type of are perfect substitutes, if $\rho \rightarrow 0$ the total labor input is Cobb-Douglas and if $\rho \rightarrow -\infty$ labor inputs are perfect complements.

From Lemma 1 we know $\frac{\partial n_a^*(z)}{\partial z} \geq 0$ and $\frac{\partial n^*(z)}{\partial z} \geq 0$. Suppose these conditions hold with strict inequality such that we can invert the optimal number of apprentices and workers. Let $n_a^{-1}(n_a)$ and $n^{-1}(n)$ denote these inverse functions. Then the firm size distribution is, $\chi(n) = \Phi(n^{-1}(n))$.

Using the market clearing conditions (4) and the wage restriction, to can find the equilibrium wages (w^*, w_a^*) and unemployment (U^*, U_a^*). For the multi-sector case, we use the more general market clearing conditions.

Calibration

To obtain the parameters of the model, we can match the firm size distribution following a method similar to Garicano, Lelarge, and Van Reenen (2016). First, we numerically compute the firm size distribution $\chi_k(n)$ for the CES case in each sector, with and without regulation. The theoretical model implies there are holes in the firm size distribution after the regulation. However, the empirical firm size distribution has notches of missing mass, but not the complete absence of firms. One way to smooth out the predictions of the model is following Garicano, Lelarge and Van Reenen by adding some noise to the observed size. Then we can derive a smoothed distribution to compare it to the data.

Additionally, we plan to use other relevant moments from our micro-data. We can match the total number of apprentices and workers by sector, before and after the policy, as well as total output, profits, and wages by sector. In the data, we observe that for every firm size there are some firms that choose to pay the fee and have zero apprentices. We can model this underlying heterogeneity to match the fraction of firms that choose to pay the fee instead of having the apprentices. All of these ideas are work in progress.

7 Final Comments

In this paper, we study the effect of apprenticeships programs on firms' outcomes. We use rich administrative data on Colombian manufacturing firms to study a policy change that shows a strong firm responses. We show these responses varied across sectors with different skill requirements. Firms in sectors with higher skill requirements avoided training apprentices. Firms in sectors with low skill requirements benefited from the policy and tried to get as many apprentices as possible. We develop a simple model to rationalize these empirical findings and quantify the welfare consequences of the policy. We are currently working on the calibration of the model to complete our quantitative exercises.

Appendix

A Firm Size Distribution by Year

Figure A.1 shows the firm size distribution is smooth for pre-reform years (1996,1998,2000 and 2002). Although the plots are more coarse given there are less observations, there aren't any visible bunching patterns around the thresholds. In contrast, if we look at the post reform years (Figure A.2) we see evidence of bunching below and at the threshold. This behavior is similar across years.

Figure A.1: Firm Size Distribution Pre-Reform

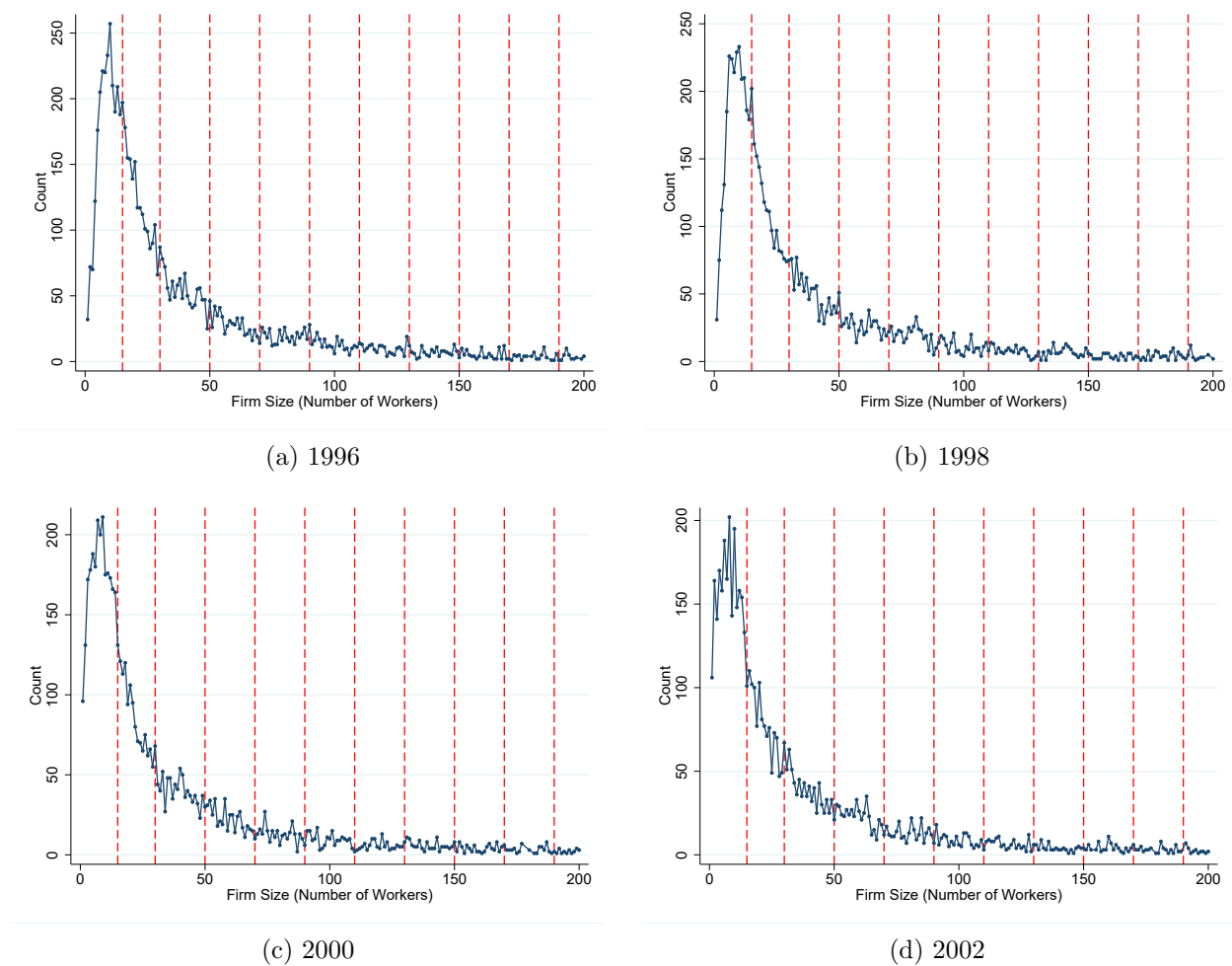
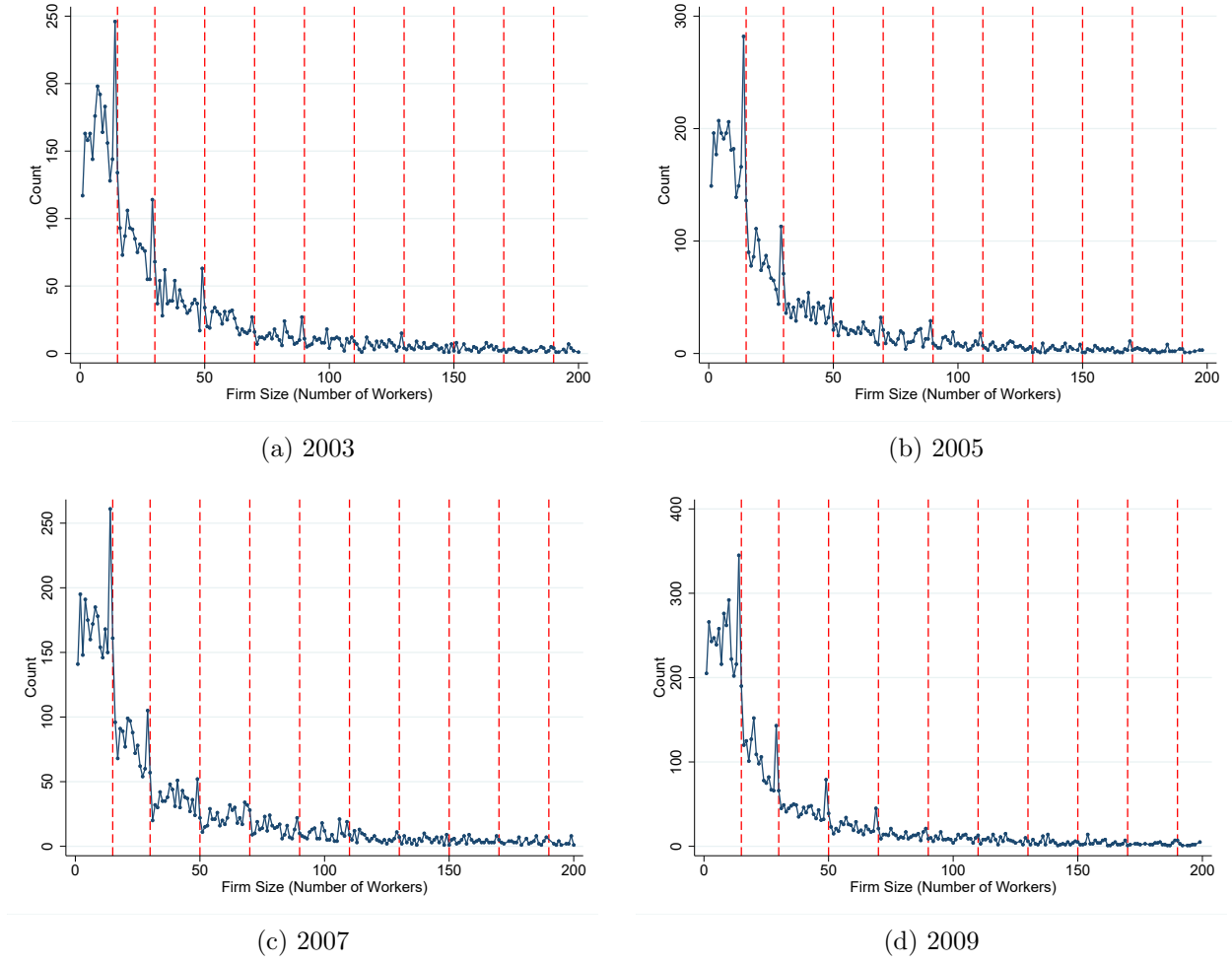


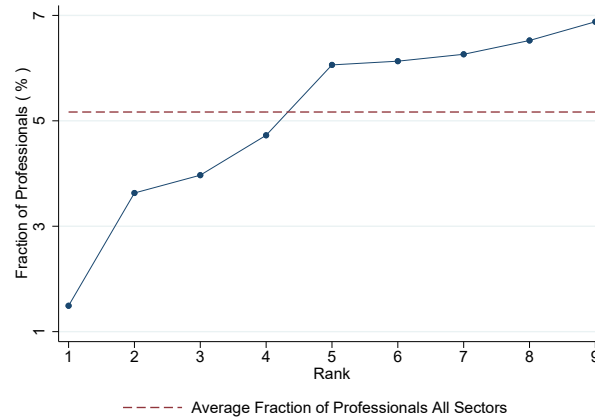
Figure A.2: Firm Size Distribution Post-Reform



B Skilled and Unskilled Sectors

Figure B.1 shows the fraction of professionals for the ranked manufacturing sectors. The dotted line represents the average fraction of professionals over total workers for all sectors. We defined the sectors below this mean unskilled sectors and sectors above, skilled sectors. Using this definition the order ranking of sectors is: four unskilled sectors (Wood Products, Textiles, Food and Beverage, Mineral Non-Metallic Products) and five skilled sectors (Paper and Editorial, Other Manufacturing, Machinery and Equipment, Metallic Products, Chemical Products).

Figure B.1: Fraction of Professional Workers by Sector



Source: Data Prepared for Ministry of Labor by National Apprenticeship Service (SENA).

References

- Daron Acemoglu and Jörn-Steffen Pischke. Why do firms train? theory and evidence. *The Quarterly Journal of Economics*, 113(1):79–119, 1998.
- Daron Acemoglu and Jörn-Steffen Pischke. The structure of wages and investment in general training. *Journal of political economy*, 107(3):539–572, 1999.
- Gary S Becker. Human capital theory. *Columbia, New York*, 1964, 1964.
- Chiara Cavaglia, Sandra McNally, and Guglielmo Ventura. Apprenticeships for young people in england: Is there a payoff? 2017.
- María Victoria Fazio, Raquel Fernández-Coto, and Laura Ripani. Apprenticeships for the xxi century: A model for latin america and the caribbean? 2016.
- Josef Fersterer and Rudolf Winter-Ebmer. Are austrian returns to education falling over time? *Labour economics*, 10(1):73–89, 2003.
- Josef Fersterer, Jörn-Steffen Pischke, and Rudolf Winter-Ebmer. Returns to apprenticeship training in austria: Evidence from failed firms. *Scandinavian journal of economics*, 110(4):733–753, 2008.
- Drew Fudenberg and Luis Rayo. Training and effort dynamics in apprenticeship. 2017.
- Luis Garicano and Luis Rayo. Relational knowledge transfers. *American Economic Review*, 107(9):2695–2730, 2017.
- Luis Garicano, Claire Lelarge, and John Van Reenen. Firm size distortions and the productivity distribution: Evidence from france. *American Economic Review*, 106(11):3439–79, 2016.
- Kathrin Göggel and Thomas Zwick. Heterogeneous wage effects of apprenticeship training. *The Scandinavian Journal of Economics*, 114(3):756–779, 2012.

- Alan B Krueger and Jorn-Steffen Pischke. A comparative analysis of east and west german labor markets: before and after unification. In *Differences and changes in wage structures*, pages 405–446. University of Chicago Press, 1995.
- Robert E Lucas. On the size distribution of business firms. *The Bell Journal of Economics*, pages 508–523, 1978.
- Rafael Novella and Yocelyn Samantha Pérez-Dávila. Are apprenticeships programs effective?: Lessons for latin america and the caribbean. 2017.
- Carlos Ospino. The effects of being subject to the colombian apprenticeship contract on manufacturing firm performance. *Documentos de Trabajo del CEDLAS*, 2018.
- OECD Economic Outlook. Recent labour market developments and prospects, 1999.
- Emmanuel Saez. Do taxpayers bunch at kink points? *American economic Journal: economic policy*, 2(3):180–212, 2010.
- David Soskice. Reconciling markets and institutions: The german apprenticeship system. In *Training and the private sector: International comparisons*, pages 25–60. University of Chicago Press, 1994.
- Stefan C Wolter and Paul Ryan. Apprenticeship. In *Handbook of the Economics of Education*, volume 3, pages 521–576. Elsevier, 2011.
- Klaus F Zimmermann, Costanza Biavaschi, Werner Eichhorst, Corrado Giulietti, Michael J Kendzia, Alexander Muravyev, Janneke Pieters, Núria Rodríguez-Planas, Ricarda Schmidl, et al. Youth unemployment and vocational training. *Foundations and Trends® in Microeconomics*, 9(1–2):1–157, 2013.