What Is the Price for Opportunity? The Effects of Employer Learning on Worker Promotions and Turnover^{*}

Sean Wang MIT

November 20, 2021

Abstract

This paper identifies the effects of firms on the career advancement of blue-collar workers and interprets these effects through the mechanism of employer learning. I use administrative data on the universe of Brazilian formal employment to study vertical promotions from production jobs to supervisory jobs, which are an important source of wage growth for most young workers. By comparing workers around job-to-job transitions, I show that differences in average firm promotion rates reflect persistent differences in the effects of firms on workers. Workers who move to a high promotion firm become substantially more likely than other job movers to be promoted, but they are even more likely to leave formal employment altogether. Correspondingly, their average long-term wage gains are negligible. I explain these effects using a model where firms differ in the rate they learn about the abilities of employed workers. High learning firms improve the efficiency of matching between workers and jobs, but these firms also exacerbate the adverse selection of unemployed workers and increase occupational wage inequality. By quantifying the parameters of the model using my estimated effects, I show that skill misallocation remains high and ex-post market power for employers can be large.

^{*}MIT Economics Department, job market paper. Contact email: swang1@mit.edu. I am grateful to Daron Acemoglu, David Autor, and John Van Reenen for their enduring guidance and support. I also thank Jonathan Cohen, Viola Corradini, Mayara Felix, Peter Ganong, Robert Gibbons, Jonathan Gruber, Ahmet Gulet, Simon Jäger, Danielle Li, Michael Piore, Charles Rafkin, Tobias Salz, Garima Sharma, Carolyn Sharzer, Carolyn Stein, Pedro Bessone Tepedino, Joonas Tuhkuri, Martina Uccioli, Michael Wong, Samuel Young, and participants at the MIT Labor Lunch and Organizational Economics Lunch for helpful comments and discussions. Access to Brazil's RAIS database is governed by the Data Use Agreement between MIT and Brazil's Ministry of Labor. I thank David Atkin and Mayara Felix for procuring MIT's access to the database, and Mayara Felix for de-identifying, harmonizing, and translating the RAIS datasets pursuant to MIT COUHES guidelines. All unmodeled errors are my own. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1745302 and the Washington Center for Equitable Growth.

1 Introduction

Research has consistently shown that where you work matters. Across a variety of countries and time periods, the same worker can expect to earn considerably different wages at different employers (Card et al., 2018). Employers, however, influence many aspects of the employment relationship besides the *level* of wages, including whether a worker is given the opportunity to advance to more complex and higher paying jobs. In Brazil, the focus of this paper, the average probability that a production worker is promoted to supervisor is nearly zero at most firms, but approximately two percent per year at the top quartile of firms.

Evidence from my data confirms that promotions are a key channel for worker wage growth and skill accumulation. Direct promotions from production worker to supervisor are accompanied by persistent earnings increases equivalent to the returns to two years of schooling. These promotions also explain five to ten percent of the lifecycle wage profile for young, blue-collar workers. Understanding the employer's role in creating promotion opportunities, especially for low skill workers, is thus critical to understanding the contributions of firms to economic mobility and inequality.

This paper uses administrative data on the universe of formal employment in Brazil to identify and interpret the effects of firms on the career advancement of blue-collar workers. I show that the unique structure of the occupational data in Brazil allows me to observe direct, vertical promotions for 70% of the formal labor market. By using flexible panel-based identification strategies, I find that the most upwardly mobile quartile of firms have persistent effects on the positive outcome of worker promotion as well as the negative outcome of formal labor force exit. I argue that these new facts are most consistent with the interpretation that firms systematically differ in their rate of learning about worker ability. I also show theoretically and empirically that this mechanism has meaningful implications for the equilibrium wage structure and the implied degrees of skill misallocation and employer ex-post market power.

The direct measurement of worker promotions using the Brazilian administrative linked employeremployee data is a key basis of this paper. The data are unique in that they clearly distinguish within occupational groups between workers who are focused on production and advanced workers who have supervisory tasks. As a result, I can directly observe when workers are promoted from production jobs to directly related and more advanced supervisor jobs. These promotions are applicable for the majority of the Brazilian workforce, including almost all blue-collar workers.

Employers are initially uncertain about worker ability, and labor markets are frictional. I establish a conceptual framework for the paper by incorporating these two features into a stylized model of job assignment and highlighting a key testable prediction about the overall effects of employers: under employer learning, firms that promote more often also fire more often. The model combines a standard learning and job assignment problem (as in Waldman, 1984; Gibbons and Waldman, 1999) with a frictional labor market where asymmetric information between employers results in the adverse selection of job movers (as in Greenwald, 1986; Acemoglu and Pischke, 1998). In the model, output is complementary in a worker's unidimensional ability and their job's complexity. It is efficient to assign high ability workers to complex jobs, keep workers of unknown ability in simple jobs, and fire low ability workers. I assume that employers are more likely to learn about a worker's ability when the worker is assigned to the complex job, so high learning employers are employers that are more willing or able to try out workers of unknown ability in complex jobs. Crucially, high learning employers are more likely to promote and fire workers because they are more active in acquiring information about the abilities of their workers.

I test the predictions of the employer learning model by estimating the effects of high promotion firms on long-term worker outcomes, and my results robustly confirm the model's predictions. High promotion firms increase the promotion probability for workers who join the firm by an additional 1.2 percentage points relative to an average baseline promotion rate of 1.7 percentage points for similar workers who join other firms. The effects decay only slightly over time and are persistent for at least seven years after moving, including across subsequent moves to other firms. On the other hand, high promotion firms also reduce the long-run formal sector employment rate for their workers by an additional 1.9 percentage points, and the effects are similarly persistent, so workers are even more likely to leave formal employment than to become promoted as a result of joining a high promotion firm. Finally, although workers who are promoted experience persistent increases in earnings, the average earnings for workers who moved to high promotion firms are comparable to average earnings for workers who moved to other firms. Taken together, my results suggest that most workers do not benefit on net from moving to a high promotion firm.

There are two key threats to identifying the effects of employers on workers: the composition of workers may differ across firms, and firms may affect workers through transient firm shocks in addition to systematic practices. I use a two-step estimation strategy to address both concerns. I first define high promotion firms as firms in the top quartile of (composition-adjusted) promotion rates for blue-collar workers using the first four years of my data. I then use the subsequent years of my data to estimate the effects of high promotion firms by comparing workers who move to high promotion firms to workers who move to other, low promotion firms. My core identification assumption is that, conditional on moving, the identity of the worker's destination firm is uncorrelated with idiosyncratic changes in the worker's labor market outcomes. This identifying assumption is weaker than the standard assumptions for estimating firm wage effects primarily because I focus on estimating longterm effects, which include any effects stemming from workers' subsequent mobility. As a result, I do not require that all mobility decisions are as good as random; instead, I assume that the job movers are comparable to each other at the time of the move, which is supported by several falsification exercises.

I also validate my baseline research design using two additional sources of aggregate variation that shift workers' job choices. The first approach follows Gibbons and Katz (1992) by using mass layoff events to purge potential biases arising from a worker's potential selection into moving employers. The second approach follows Oreopoulos et al. (2012) in spirit by instrumenting for the worker's destination firm with the local hiring share from all high promotion firms (excluding the worker's destination firm) to purge potential biases arising from the worker sorting based on idiosyncratic shocks or unobserved trends. My estimates remain similar when using these aggregate variation for identification, which supports the interpretation that the estimated effects are driven by the effects of firms rather than by worker sorting.

I then use a series of additional empirical exercises to show that the estimated effects are consistent with employer learning and inconsistent with several alternative explanations. Differences in promotions between firms appear to reflect real differences in job assignments and earnings rather than simply differences in the willingness to label otherwise identical workers as supervisors. Similarly, worker exits from formal employment are explained by negative outcomes like layoffs or firings rather than voluntary quits. The negative employment outcomes are also concentrated on workers who are likely to be promoted, which is consistent with learning and inconsistent with greater volatility in firm labor demand. Finally, corroborating survey evidence from managers across a variety of countries supports the interpretation that the effects reflect systematically different firm responses to general economic mechanisms rather than any Brazil-specific institutional feature.

After establishing that the direct employer effects are consistent with employer learning, I then explore the effects of high learning employers on *all* workers in their market in addition to direct effects on their employed workers. I theoretically clarify this indirect mechanism by endogenizing the labor market parameters of my model through an initial vacancy creation stage, and I empirically test the resulting predictions about the overall wage structure. High learning firms are more likely to fire low ability workers, which exacerbates the adverse selection of workers who change jobs. As a result, the secondary market is more pessimistic about the expected quality of incoming workers and makes lower wage offers. In equilibrium, wage differentials between promoted and non-promoted workers rise since non-promoted workers are more exposed to the softer secondary market competition. These additional predictions about occupational wage inequality are also robustly supported by the data. Moving from the 10th to the 90th percentile of municipalities (in terms of the local employment share of high promotion firms) increases the local wage premium for supervisors by 39% of its average value.

The model also establishes a framework for quantifying the economic magnitude of the learning mechanism using my estimated employer effects. Although I make strong assumptions on wage setting to close the model in an analytically tractable manner, the basic components of the model that govern job assignment and turnover can be mapped to the estimated effects under a wide range of realistic wage setting mechanisms, including bargaining and binding wage floors. The key restriction for model identification is that workers in the analysis sample do not systematically differ based on their destination firms. This restriction is stronger than the identification assumption used to estimate the employer effects, but it is also consistent with that identification assumption and supported by the data.

Quantifying the model shows that the rate of employer learning is low on average, including at high promotion firms. As a result, skill misallocation is high -87% of workers who would be suitable for supervisory occupations end up in elementary occupations or outside of formal employment. Nevertheless, the rate of employer learning is sizable relative to idiosyncratic turnover, so the scope for employer ex-post market power is also substantial. I estimate that job movers are only 71% as likely to be of high ability as the general population, which suppresses outside options for workers and amplifies the value of private information held by employers.

Related Literature

This paper primarily contributes to the literatures on firm heterogeneity in labor market outcomes, worker dynamics within the firm, and firm promotion policies. To my knowledge, it is one of a few papers to focus on occupational outcomes using economy-wide administrative data, and it is the first of these papers to document the firm-level link between promotion opportunity and separation risk. It is also part of a series of papers that combines quasi-experimental estimates of firm heterogeneity with a theoretical framework for quantifying their economic implications, and it is the first of these papers to focus on the implications of employer learning.

A well-established literature has documented that differences between firms pass through to

workers. A series of papers use linked employer-employee data and an exogenous movers research design to estimate the dispersion in the level of firm wage premia (for example, Abowd et al., 1999; Card et al., 2013, 2016; Song et al., 2019), including in the Brazilian context (Alvarez et al., 2018; Gerard et al., 2018). Related papers using worker data have also found that firms experiencing shocks tend to share a portion of the shocks with workers (Kline et al., 2019; Lamadon et al., 2019). There are relatively fewer papers focusing on firm differences in long-run effects, but the papers that do so tend to focus on interpreting long-term wage changes through search, human capital accumulation, or a mixture of the two (Bagger et al., 2014; Herkenhoff et al., 2018; Gregory, 2019; Jarosch et al., 2019; Arellano-Bover, 2020; Taber and Vejlin, 2020; Addario et al., 2021).¹ Within this literature, Jarosch (2015) shows that firms differ in layoff risk that erodes workers' long-term earnings, and Arellano-Bover and Saltiel (2020) shows that firms differ in wage growth that may be correlated with occupational growth. My paper supports the idea that these two patterns reflect firms' causal effects rather than worker sorting. Moreover, I argue that the two outcomes are linked through the mechanism of employer learning.²

A similarly rich literature has analyzed the firm's role in changing a worker's skill mix. One prominent explanation for a worker's rising task complexity over time is that firms provide either direct or indirect training (Becker, 1964; Mincer, 1974; Acemoglu and Pischke, 1998; Jovanovic and Nyarko, 1997; Lazear, 2009). An alternative explanation is that information about a worker's productivity is revealed over time, so tenure profiles reflect selection rather than investment (Jovanovic, 1979; Waldman, 1984; O'Flaherty and Siow, 1995; Farber and Gibbons, 1996; Altonji and Pierret, 2001; Golan, 2005; Lange, 2007). Of course, the two explanations are not mutually exclusive and may interact (Gibbons and Waldman, 1999; Autor, 2001; Kahn and Lange, 2014; Pastorino, 2019). This paper shows that learning is relevant even for older production workers and can rationalize the observable patterns in job assignments and turnover. However, my data are not suited to directly study training, and I do not reject the role of either direct training or indirect training (which may be a conduit for learning or promotions). In fact, I find evidence in support of asymmetric information as a meaningful source of labor market frictions, which would imply that firms also benefit from paying for general training.

Finally, there is substantial interest in describing and interpreting firms' promotions decisions. Case studies have documented that firms commonly draw higher level workers from their pool of lower level workers rather than from external sources (Doeringer and Piore, 1971; Baker et al., 1994). Recent studies on white-collar workers also show that, although firms attempt to target workers for promotion, their decisions are not always efficient or fair (Benson et al., 2019; Cullen and Perez-Truglia, 2021). Particularly related is Friedrich (2020), which uses administrative Danish data to show that more productive firms are more likely to use internal labor markets for filling top and middle managerial positions. My conclusions are consistent with Friedrich (2020) – actively promoting firms

¹Note, though, that interpreting wage growth as human capital typically requires making strong assumptions about wage setting, since many mechanisms like job search and dynamic contracting would also generate wage growth without any changes in worker productivity. For example, workers who continue to search for new jobs while employed may receive wage increases whenever they receive a competing outside offer (Postel–Vinay and Robin, 2002; Caldwell and Harmon, 2019). Models of optimal dynamic contracting often feature increasing wages over a worker's tenure even if the worker's productivity is constant or decreasing (Lazear, 1979; Burdett and Coles, 2003).

²Another portion of the literature seeks to answer whether worker *utility* is equalized across firms and concludes that non-wage characteristics are dispersed across firms (Sorkin, 2018; Maestas et al., 2018). My paper supports those conclusions by showing that career advancement potential is a form of indirect compensation that differs across firms. I also raise the possibility that workers may be uncertain about employers' non-wage qualities at the time of hire.

improve matching efficiency at the cost of higher adverse selection and wage inequality. I additionally provide quasi-experimental evidence supporting the causal effects of firms, and I show that these mechanisms are also relevant for the majority of the blue-collar workforce and have long-term consequences on workers' labor force attachment.

The rest of the paper proceeds as follows. Section 2 establishes a conceptual framework by constructing a model of job assignment and deriving predictions about employers' effects. Section 3 describes the Brazilian institutional setting and the administrative data. Section 4 discusses the identification assumptions for estimating employers' effects on workers. Section 5 reports estimates of the effects of high promotion firms. Section 6 rules out potential alternative mechanisms using a series of additional checks. Section 7 then extends the baseline model to characterize the additional equilibrium effects of high learning firms. Section 8 discusses the structural quantification approach and reports the model's estimates. Section 9 concludes.

2 Model of promotions and exit under employer learning

To fix ideas about the possible effects of high promotion firms, I begin by characterizing a stylized model of job assignment. I incorporate three critical features that are realistic for my setting. First, firms learn about the quality of their employed workers over time. Second, information about worker quality is asymmetric between firms. Third, labor market matching is frictional and random. The resulting model generates a key prediction – firms that are more likely to promote workers are also more likely to fire workers – which I test in Section 5 by identifying the effects of firms on worker promotions and exit. The model in this section also serves as a basis for exploring the equilibrium effects of employer learning on the regional wage structure in Section 7 and for quantitatively interpreting the estimated employer effects as structural parameters in Section 8.

2.1 Basic setup

I follow the framework of Waldman (1984); Gibbons and Waldman (1999) in modeling vertical job ladders as the optimal matching between workers' unidimensional ability and jobs' returns to ability. I assume that workers are one of two types: high ability ($\theta = \theta_H$) with probability α , and low ability ($\theta = \theta_L$) with probability $1 - \alpha$. There are two possible job assignments $j \in \{1, 2\}$ that are supermodular in worker ability and crossing:

$$f_{2}(\theta_{L}) < f_{1}(\theta_{L}) < 0 < f_{1}(\theta_{H}) < f_{2}(\theta_{H}).$$

Figure 1 shows an example of the expected output from each of the two occupations as a function of the probability that the worker is of high ability. The complex job (f_2) has higher returns to ability but lower output for low ability workers. As a result, expected output is maximized in the high complexity job (j = 2) if the worker is likely to be of high ability and in the low complexity job (j = 1) if the worker is likely to be of low ability. Moreover, I assume that $f_j(\theta_L) < 0$ for $j \in \{1, 2\}$, so workers who are revealed to be of low ability are not productive in either job, which introduces a motive for firms to fire workers. Finally, I assume that

$$\alpha f_1(\theta_H) + (1 - \alpha) f_1(\theta_L) > \alpha f_2(\theta_H) + (1 - \alpha) f_2(\theta_L),$$

so it is output maximizing to assign unknown workers to the low complexity occupation and the promotions problem is not trivial.

The timing of worker and firm actions and the realization of events are summarized in Figure 2. Workers are initially employed at either a high learning firm (f = H) with probability ρ or a low learning firm (f = L) with probability $1 - \rho$. At the start of employment, each firm f randomly assigns each worker to a trial in the complex job (f_2) with probability Q_f or a trial in the simple job (f_1) with probability $1 - Q_f$. The firm then observes the worker's output from the trial and updates its beliefs about the worker's ability. The complex job always reveals the worker's ability and the simple job never reveals the worker's ability. So, a high learning firm is more likely to learn about the ability of each worker because it is more likely to assign them to the more complex job (i.e., $Q_H > Q_L$).³

Based on its updated beliefs, the firm then decides either to fire the worker or to offer them a wage and job assignment. A fraction δ of workers exogenously separates from their employers. The remaining workers choose to either accept the firm's offer or leave the firm. All workers who separate from their initial employers encounter the secondary market with probability g.

Firms in the secondary market compete by making wage offers to workers. These firms have access to the same set of production technologies as incumbent firms, but they do not observe the ability of any worker, so there is asymmetric information. If a worker was offered a promotion at their previous employer, there is a probability κ that they can convince the secondary market of this fact. Firms do not observe any other information about incoming workers, so they do not know whether the worker was fired by their previous employer or they voluntarily quit.

At the end of the period, each firm employing a worker of type θ in job j with wage w receives $\pi = f_j(\theta) - w$, while the worker receives w. Workers and firms are risk-neutral, and their outside options are normalized to 0.

2.2 Partial equilibrium and direct effects

The partial equilibrium of the model takes the share of high learning firms (ρ) and the secondary market contact rate (g) as given and assumes that workers and firms maximize their expected wages and profits. Since the game is one of imperfect information, I use the Perfect Bayesian equilibrium as my solution concept. The Perfect Bayesian equilibrium in this model is a set of worker and firm strategies such that the following conditions are satisfied:

- 1. Workers make turnover decisions that maximize expected wages given the incumbent firm's wage and job assignment and the expected secondary market wage offers
- 2. Incumbent firms make wage and job offers that maximize expected profits given the worker's turnover decision and the expected secondary market wage offers
- 3. Firms in the secondary market make wage offers that maximize expected profits given their beliefs about the expected ability of each worker in the secondary market

³An equivalent formulation, which generates the same predictions, is to assume that the firm simply learns about the ability of each worker with some probability Q_f before making job and wage offers. A high learning firm is then a firm that learns about the worker's ability more often (so $Q_H > Q_L$). I use the "trial promotion" formulation because it is more consistent with the dynamics of my estimated promotion effects, as discussed in Section 5. It is also equivalent to assume instead that both jobs are partially informative as long as the complex job is more informative. I make the more stark assumption here to simplify notation.

4. Firms in the secondary market have rational beliefs about the ability of each worker given the turnover decisions of workers and the wage and job assignment decisions of incumbent firms

I consider the primary and economically interesting equilibrium in which an incumbent firm promotes workers who are revealed to be of high ability, fires workers who are revealed to be of low ability, and retains workers whose abilities remain unknown. Working backward, the secondary market is perfectly competitive, so firms offer the expected output for each worker and make zero profits in equilibrium. Any worker that successfully convinces the secondary market that they were previously promoted must have been revealed to be of high ability. Meanwhile, other workers in the secondary market are a combination of high ability workers (who exogenously separated and were not recognized as high ability by the secondary market) or low ability workers (who either exogenously separated or were fired). The probability that an unknown secondary market worker is of high ability is then

$$\alpha' = \frac{\alpha \delta \left(1 - \bar{Q}\kappa\right)}{\alpha \delta \left(1 - \bar{Q}\kappa\right) + (1 - \alpha) \left(\delta + (1 - \delta)\bar{Q}\right)},\tag{1}$$

where $\bar{Q} = \rho Q_H + (1 - \rho) Q_L$ is the average rate of learning in the economy. These workers are adversely selected ($\alpha' < \alpha$) both because high ability workers are more likely to enter the secondary market as promoted workers and because low ability workers are also more likely to be fired. Secondary market wages w_j^S for each job j are

$$w_1^S = \underbrace{\alpha' f_1(\theta_H) + (1 - \alpha') f_1(\theta_L)}_{\equiv E[f_1(\theta)|\alpha']}$$

$$(2)$$

$$w_2^S = f_2(\theta_H)$$

where I assume that α' is sufficiently high so that the adversely selected workers are still expected to be productive:

$$E\left[f_1\left(\theta\right)|\alpha'\right] > 0. \tag{3}$$

An incumbent firm makes take-it-or-leave-it offers to workers, so it is sufficient to offer the expected outside option for the worker to accept. I assume that job take-up decisions are made before the worker makes contact with the secondary market (i.e., there is no on-the-job search), so secondary market wages are discounted by the probability that the worker encounters the secondary market (g). Promoted workers run the additional risk that they may not successfully convince the secondary market that they were previously promoted (i.e., $\kappa < 1$). So, optimal incumbent wage offers w_i^I for each job j are

$$w_{1}^{I} = g w_{1}^{S} w_{2}^{I} = g \left[\kappa w_{2}^{S} + (1 - \kappa) w_{1}^{S} \right].$$
(4)

Notice that even in the case when g = 1, so there are no re-employment frictions, the incumbent employer still earns positive profits from each unpromoted worker due to adverse selection in the secondary market, since $w_1^S = E[f_1(\theta) |\alpha'] < E[f_1(\theta) |\alpha]$. Meanwhile, w_2^S reflects the true productivity of promoted workers, so the imperfect transmission of information about promoted workers ($\kappa < 1$) ensures that an incumbent firm also retain some informational rents from promoted workers.

For the conjectured job assignments to be optimal for an incumbent firm, the following conditions

need to hold

$$E\left[f_1\left(\theta\right)|\alpha\right] - w_1^I \ge 0$$
$$f_2\left(\theta_H\right) - w_2^I \ge f_1\left(\theta_H\right) - w_1^I$$

The first condition ensures that the firm will find it profitable to retain workers whose abilities remain unknown by offering them the low complexity job with wage w_1^I .⁴ This condition is implied by my assumption that employment in the secondary market is viable for the adversely selected workers (Equation 3). The second condition is an incentive compatibility condition that ensures the firm will find it more profitable to promote high ability workers and pay the higher wage w_2^I than to keep them in the low complexity occupation and pay the lower wage w_1^I .

The conditions for the existence and uniqueness of the equilibrium and the characterization of the equilibrium strategies are summarized in the following Proposition 1, and a detailed proof of the proposition is in Appendix B.

Proposition 1. If information about job assignments on the secondary market is sufficiently weak (so κ is sufficiently small), then a unique Perfect Bayesian equilibrium exists. In this equilibrium, (i) job assignments for workers are efficient (given firms' information about workers) (ii) all turnover is involuntary (iii) wages are given by Equations 2 and 4.

To see why the degree of asymmetric information (κ) is key to satisfying the conditions for existence, note that a low κ relaxes the key constraints for both the secondary market and the incumbent firm. A low κ ensures that promoted high ability workers may nevertheless enter the secondary market in the same pool as fired low ability workers, which helps offset the degree of adverse selection in the market. Meanwhile, a low κ also softens wage competition for promoted workers at the incumbent firm, since workers are then less likely to maintain their higher position if they go to the secondary market. This reduces the wage increase necessary to retain a worker upon promotion.

The partial equilibrium setup is sufficient for comparing the outcomes of otherwise identical workers who were initially matched to a high learning firm $(Q_f = Q_H)$ as opposed to a low learning firm $(Q_f = Q_L)$. Since workers may be of either high ability or low ability, being matched to a high learning firm introduces greater opportunities for promotion, but also greater risk of becoming unemployed whenever the probability of re-employment upon separation (g) is less than 1. Proposition 2 formalizes this comparison, and the accompanying proof is also in Appendix B.

Proposition 2. In the equilibrium described in Proposition 1, workers initially employed at high learning firms are (i) more likely to be promoted and (ii) more likely to become unemployed than workers initially employed at low learning firms.

Proposition 2 makes the key predictions about the direct effects of firms that I will empirically test in Section 5. Specifically, the testable predictions are that in the presence of systematic differences in employer learning:

1. Some firms will be systematically more likely to promote workers

 $^{^{4}}$ The firm will never find it optimal to offer those workers the high complexity job since under these assumptions, their expected output is lower but the required wages to retain the worker are higher.

2. The firms that are more likely to promote workers are also more likely to fire workers

The model assumes that labor market matching is random, so the workers initially working at high learning firms are comparable to the workers initially working at low learning firms. However, the workers that exit firms are not comparable between high and low learning firms due to firings. In the presence of this selection, simply comparing the cross sectional differences between workers at different firms would overstate the causal effect of high learning firms on worker promotions and understate the causal effect on worker exits. Correspondingly, I focus on testing the model's predictions by estimating the effect of firms on newly hired workers, and I ensure that the effects are not driven by differences in worker composition.

3 Setting and data

The empirical setting is the Brazilian formal employment sector between 2003 and 2015, which is particularly well suited to examining the effects of employers on workers' careers for three reasons. First, Brazil's uniquely detailed administrative data allow researchers to follow approximately 70% of the formally employed workforce along direct lines of progression in their occupation groups. Second, although Brazilian labor market institutions have rigid components, employers generally have the flexibility to assign workers to different jobs. Finally, the rate of formal higher education in Brazil is low compared to high income countries, so employers can be expected to play a larger role in human capital accumulation or signaling.

3.1 Background on the Brazilian labor market

Brazil's labor market environment and institutions have been the subject of extensive research. Instead of trying to write a complete account of the Brazilian labor market, I focus on two aspects that are particularly important for the interpretation of my empirical strategy and results. I also briefly summarize several other details that provide additional context.

First, Brazil, like many other developing countries, has a sizable informal sector of the labor market. In Brazil, informal jobs do not have a signed "work card" (Carteira de Trabalho) and consequently are not subject to taxes and labor market regulations. Estimates of the size of the informal sector can vary since informal jobs are by definition missing from official registers. Recent estimates range as high as approximately 50% in metropolitan areas to as low as approximately 20% for all prime age workers; these estimates are comparable to informality rates in other developing economies. Firms in the formal labor market are generally more productive, pay higher wages, employ a greater share of educated workers, and are required to provide legally mandated employment protections and unemployment insurance for their workers.⁵ In my data, I can observe whether a worker leaves the formal labor market, but not whether the worker enters the informal labor market versus unemployment. I generally interpret these exits as negative employment outcomes for the worker (and provide supporting evidence for this in Section 6.2), but it is important to note that I do not equate leaving the formal labor market as necessarily reflecting unemployment.

 $^{{}^{5}}$ See, for example, the summary in Perry et al. (2007) and estimates in Gerard et al. (2018); Haanwinckel and Soares (2020); Dix-Carneiro et al. (2021).

Second, although the Brazilian labor code mandates some employment protections for workers in the formal sector, firms have the latitude to determine both the size and composition of their workforce. Legal protections mandating job security for workers were largely eliminated for employment contracts after 1966, and firms are allowed to dismiss workers without cause as long as they provide notice and severance penalties. The length of notice is typically 30 days (and can be as high as 90 days for workers with long periods of service), and the severance payment is up to 4% of the worker's total earnings while employed at the firm.⁶ Firms are exempt from severance payments for voluntary quits or dismissals with cause (for more details, see Gonzaga et al. 2003; OECD-IDB 2014). By the OECD summary index of employment protections for workers with regular contracts, Brazil is scored as more flexible than countries like Denmark, the United Kingdom, and New Zealand. A key aspect of my interpretation is that some firms are more likely to layoff or fire workers after learning new information about the worker, and the legal flexibility afforded to employers is consistent with this interpretation.

Several other factors about the labor market in Brazil provide helpful context but are less central to the interpretation of my results. Labor unions are prominent in Brazil and bargain at the sectoral and the firm level. Firm-level collective bargaining agreements tend to cover all workers at an establishment (rather than varying by occupations or union membership) due to the Brazilian practice of universal coverage. The Brazilian labor code also explicitly prohibits nominal wage reductions, except those that are negotiated through collective bargaining (Lagos, 2019). The rate of tertiary education is generally low but has been experiencing rapid growth from 10% for 25-34 year olds in 2007 to 17% in 2017 (OECD, 2019). Finally, the macroeconomic environment around my period of study is generally stable. Brazil's period of hyperinflation ended with the introduction of the real in 1994, and inflation during my period of study ranged between 4-9%. While the period that I study straddles the Great Recession, its effects were muted in Brazil; instead, the country experienced a larger and more persistent recession starting in 2014.

3.2 Data and sample restrictions

My data on worker outcomes and firm characteristics come from the Brazilian Ministry of Labor's Relaçao Anual de Informações Sociais (RAIS), a worker-level dataset containing reporting data on all formal employment contracts. The data are likely to be fairly complete and high quality since the government mandates reporting to RAIS for all formal sector employers and penalizes late or missing filings. The dataset has also been used in several recent studies about the Brazilian labor market (including Menezes-Filho et al., 2008; Dix-Carneiro and Kovak, 2017; Alvarez et al., 2018; Gerard et al., 2018; Dix-Carneiro et al., 2021).

I observe key information about each formal employment contract, including its occupation, duration, contracted hours, contracted salary, and average monthly earnings. Furthermore, I observe detailed demographic information about the worker, including their gender, age, educational attainment, and race, as well as basic information about the employer's industry and location. The records are linked over time by longitudinal identifiers, so I can follow workers over time across different firms. The RAIS separately records each employment contract, and a worker may hold several jobs over a year due to either job transitions or multiple part-time jobs. I construct an annual panel of

 $^{^{6}}$ In the case of dismissals without cause, 80% of the penalty is paid out to workers and 20% is added to the state unemployment insurance fund

worker employment histories from the collection of all contracts by selecting the long-term employment contract with the highest average earnings for each worker.⁷ Additional details about the data construction are in Appendix C.1.

I consider outcomes between 2003, the first year when worker data are reported under the revised Brazilian occupational codes, and 2015, the last year of my data. For most worker-level analyses (unless otherwise specified), I restrict my sample to prime age workers who are between the ages of 25 and 50, since they have a high attachment to the formal labor force and are less likely to experience formal schooling or retirement. I discuss any additional sample restrictions for the implementation of my identification approaches in the corresponding parts of Section 4.

3.3 Measuring promotions

A distinctive feature of the Brazilian data is that I can observe direct lines of progression from worker to supervisor for most blue-collar occupational groups. Worker occupations are reported under the Classificação Brasileira de Ocupações (CBO) system, which is similar to the International Standard Classification of Occupations (ISCO) in that it groups jobs into a hierarchical structure based primarily on the type and complexity of the tasks involved. However, the Brazilian system, especially after its 2002 revision (CBO-02), is particularly suitable for the measurement of worker promotions for two reasons. First, it organizes occupations into occupational groups by natural lines of progression in addition to task content, so occupations that are often related for workers but do not share narrowly-defined tasks (like loom operators and fabric dyers) are explicitly grouped. Second, it consistently distinguishes between supervisors, who are advanced production workers with additional managerial responsibilities, and more elementary line workers within each occupational group. As a result of these two features, I can interpret changes from production occupations to supervisor occupations within the same general occupational group as reflecting a direct, vertical promotion.

Table 1 shows an example of an occupational group with a directly observable line of progression. The first two digits of the occupation code indicate the main occupational group, a "0" in the third digit indicates the sub-group of supervisory occupations, whereas other codes in the third digit indicate other sub-groups that contain production occupations but do not necessarily have a clear hierarchical structure. This basic structure applies to nearly all of the non-professional occupational groups in the Brazilian data. Figure A.1 shows that around 70% of workers in the formal sector belong to a 2-digit occupational group with a clear supervisor-worker line of progression and that this share is fairly stable over the period that I study.⁸

In this paper, I define a promotion as moving from any production occupation belonging to a supervisor-track occupational group to a supervisor occupation. Figure 3 shows that promotions are reasonably common. Furthermore, around half of promotions are purely vertical moves within the same occupational group. Promotion rates are fairly stable and free of secular trends over my sample, although they did begin to taper around 2014 when Brazil experienced a recession.

Two descriptive facts support the argument that my measure of promotions captures an important source of worker growth. First, promotions are valuable. Table 2 shows the estimated wage premium

⁷Defined as a contract that covered at least 6 months out of the year and entailed at least 20 contracted hours per week.

 $^{^{8}}$ A reclassification of the CBO-02 system occurred in 2008 and slightly increased the share of workers that belong to a relevant occupational group, but it does not substantively affect my approach or results.

for supervisors relative to other production workers. The average cross-sectional wage premium is 63 (s.e. 1.8) log points. Controlling for worker characteristics accounts for 39% of this difference and controlling for unobserved differences through worker fixed effects accounts for another 34%, which implies that supervisors are positively selected. Even when I focus on within-worker changes in earnings, promotions are accompanied by wage increases of 17 (s.e. 0.35) log points, which is equivalent to the returns to an additional 1.7 years of education in Brazil.⁹

Second, promotions consistently explain part of the lifecycle wage profile for young production workers. Table 3 compares the estimated lifecycle wage profile for workers between the ages 25 and 35 before and after including the promotion measure as an explanatory variable. When controlling for only basic worker characteristics, an extra year in age increases earnings by 1.73 (s.e. 0.03) log points. Adding promotions as an explanatory variable decreases the age coefficient to 1.63 (s.e. 0.03) log points, which implies that promotions explain approximately 5% of the age profile. The estimated age coefficient falls as I add additional controls for occupational characteristics and firm wage premia, but the estimated promotion wage premium stays relatively stable. As a result, the share of the age profile that is explained by promotions increases to 11%.¹⁰ Although this exercise is descriptive, the results confirm that promotions capture a meaningful change that is distinct from other factors like schooling and employer upgrading.

4 Identifying the effects of firms

I focus on identifying the effects of *existing firms* on worker promotions and job turnover, so the ideal experiment is to randomly assign workers to various firms while taking as a given that those firms are a bundle of underlying practices. As a result, my identification strategy centers around changes in a worker's firm assignment, rather than changes in a firm's practices or changes in a worker's promotion likelihood. The baseline approach compares changes in worker outcomes among job-movers by the type of the destination firm, and two extensions relax the key identification assumptions by using aggregate shifters of workers' job choices.

My approach is similar to the two-step grouped fixed effects estimator of Bonhomme et al. (2019). I first separate firms into distinct groups using observational data, and I then estimate the effects of each group of firms on workers while allowing for flexible dynamics. However, my identification assumption is weaker than that of Bonhomme et al. (2019) since my estimand of interest is the long-term effect of differences between firms.¹¹

⁹Controlling for firm-wage premia by subtracting the estimated AKM firm wage fixed effect from worker earnings increases the estimated promotion premium from 17 log points to 20 log points. These estimates are generally larger than the average wage increases accompanying promotions in the literature, but smaller than the large jumps in pay at the top of corporate hierarchies (see, for example, Murphy, 1985; Baker et al., 1994; McCue, 1996; Blau and Devaro, 2007).

¹⁰My estimate is generally comparable to the estimate from McCue (1996) using self-reported promotions from the U.S. PSID. For comparison, Topel and Ward (1992) estimate that a third of wage growth over the first 10 years of employment in the U.S. is accounted for by wage gains at job changes, which would include promotions, and Bagger et al. (2014) estimate that human capital growth accounts for 20-25% of the life-cycle wage profile for low to medium educated workers in Denmark.

¹¹As a concrete example, suppose firm wages are static and vary only by level. In this setting, firms may have dynamic effects on worker earnings if they disproportionately lead their workers to move to high wage firms. Estimators that seek to identify firms' wage policies would need to make assumptions about workers' mobility decisions to net out the contribution of subsequent high wage firms from the dynamic effects of prior firms. On the other hand, I seek to estimate the overall effect of the prior firm, including any effects that arise from workers subsequently moving to high wage firms.

Generalizing the movers research design into a two-step event study framework has three additional benefits. First, I can explicitly separate the timing between the classification of firms and the estimation of employer effects to ensure that the effects reflect permanent differences between firms rather than different realizations of transitory firm shocks. Second, I can follow standard event study methodology in conducting falsification tests of my identifying assumptions through pre-trend and balance tests. Finally, it is straightforward to use flexible estimators that estimate the relevant average treatment effect by combining cohort-level treatment effects with explicit weights.

4.1 Defining high promotion firms

As a first step, I classify firms into two groups using firms' composition-adjusted promotion rates between 2004 and 2006. For each year between 2004 and 2006, I estimate the worker-level regression

$$Promoted_{it} = \beta_t X_{i,t-1} + \gamma_{ot} + \eta_{jt} + \varepsilon_{it}, \qquad (5)$$

where Promoted_{it} = 1 if a worker was promoted from production worker to supervisor between years t - 1 and t. $X_{i,t-1}$ and γ_{ot} adjust for differential promotion rates by the worker's observable characteristics (a quadratic in age interacted with gender and indicators for education, race, and state) and the worker's occupational group, respectively.¹² η_{jt} is the firm's residualized promotion rate in year t. I average over all three years of estimates for each firm to calculate the firm's average promotion propensity $\eta_j = E[\eta_{jt}]$. Firms with fewer than 10 promotion-track workers in at least one of those three years and firms who did not exist in all three years are unlikely to generate precise estimates of η_j or survive through to the movers analysis, so I drop them from the classification sample. I similarly exclude public sector firms, which may have different organizational structures and internal incentives.

Figure A.2 plots the distribution of the average promotion propensity η_j for the remaining sample of firms, winsorized at the 5th and 95th percentiles. The distribution is highly skewed, with a small tail of firms that are highly active in promoting workers and a majority of firms that promote few (if any) workers. Some of the dispersion is mechanical since residualizing promotion rates by worker characteristics can generate dispersion in the firms' estimated fixed effects even when the firms' promotion rates were uniformly zero. Splitting firms into two groups based on η_j then ensures that my estimates of differences between firms are not primarily driven by small differences in residualized rates between firms that did not promote any workers.

For the rest of the paper, I define high promotion firms as firms with an average promotion propensity (η_j) that is in the top quartile of all firms in this classification sample, and low promotion firms as all other firms in the sample. Table 4 shows the differences in firm characteristics between high and low promotion firms between 2003 and 2006. High promotion firms are 49 log points larger on average, so they actually constitute 34% of employment during this period. They are also higher paying and have a greater share of supervisors in the relevant occupational groups.¹³ High promotion firms are responsible for almost the entirety of promotions within this sample of firms,

 $^{^{12}}$ To facilitate interpretation, I restrict the sample to workers who are in the same occupational group in both years. However, the estimates are highly correlated if I include workers who switch occupations or restrict the sample further to include only job stayers. For more details, see Table A.1 and Appendix C.1.

¹³There may be concerns that supervisor shares are mechanically higher at high promotion firms since these firms were defined by having promoted workers. I also consider the prior year's supervisor share as a check, and the results are slightly smaller but similar.

which supports pooling the remaining firms together as one group. Finally, promotions may reflect in part the realization of positive firm shocks – high promotion firms are both faster growing and have higher wage growth for incumbent workers. This correlation between promotions and firm growth rates motivates testing for persistent effects using workers who later join these firms.

4.2Baseline research design

Estimating equation and identifying assumption

My baseline research design compares the change in outcomes between workers who move to high promotion firms and workers who move to low promotion firms. The primary identification assumption is that for a worker who is making an employer-to-employer transition, the type of the worker's destination firm is uncorrelated with idiosyncratic changes in the worker's labor market outcomes. This identification assumption is implied by standard exogenous mobility assumptions that are used in the literature for estimating firm wage effects, but it is strictly weaker. Specifically, I also allow for workers to select into moving based on idiosyncratic worker-level shocks and for firms to have persistent effects on workers after they leave.

To fix ideas, consider workers who make an employer-to-employer move in a single cohort year c. Let $H_i = 1$ if worker i moved to a high promotion firm and $H_i = 0$ if they moved to a low promotion firm. My estimation equation for the dynamic effects of the high promotion firm on worker outcome y_{it} is

$$y_{it} = \sum_{\tau} \beta_{\tau} \left(H_i \times I_t^{\tau} \right) + \alpha_i + \pi X_{it} + \theta_t B_{i,c-1} + \gamma_{sot} + \varepsilon_{it}, \tag{6}$$

where I_t^{τ} are event-time indicators that equals 1 when $t-c=\tau$ and 0 otherwise (with $\tau=-1$ as the omitted time period).¹⁴ α_i is a worker fixed effect, X_{it} is a vector of time-varying worker covariates, and $B_{i,c-1}$ is a vector of baseline worker covariates from year c-1 that may affect worker outcomes in other periods through the time varying coefficients θ_t . Finally, I saturate the regression with γ_{sot} , a set of state-by-baseline-occupational-group-by-time fixed effects that control flexibly for aggregate trends in market conditions or skill prices.¹⁵

Note that by defining the estimand of interest β_{τ} as the effect of moving to a high promotion firm after τ years, rather than the effect of staying at the high promotion firm for τ years, I do not distinguish between workers who remain at the high promotion firm from workers who subsequently leave. So, my estimate includes the direct effects of the firm on its stayers as well as the persistent effects of the firm on workers' mobility decisions and subsequent outcomes at other firms. I make this choice for two reasons. First, some models of human capital accumulation predict that workers would realize most of the gains after leaving the employer responsible for providing the human capital, and this estimand ensures that I appropriately capture these channels.¹⁶ Second, I do not need to make any assumptions about the worker's mobility decisions after the year of the initial

¹⁴In pooled specifications, I collapse event time into two periods – the near term ($\tau \in [0,2]$) and the long-term

 $^{(\}tau > 2)$. ¹⁵In the baseline specification, the time varying covariates X_{it} are quadratic trends in age that vary by gender and AKM form indicators for educational attainment, while the baseline covariate $B_{i,c-1}$ is the origin firm's estimated AKM firm effect. These controls explain any differences in baseline characteristics of the high versus low promotion firm movers, as shown in the balance test discussed later in this section.

 $^{^{16}}$ For example, employers that provide training may demand compensating differentials through lower wages (Becker, 1964), or employers may specialize in jobs that are "stepping stones" for more complex jobs at other firms (Jovanovic and Nyarko, 1997).

move, which itself may be a result of the worker's destination firm.

The key assumption for the identification of β_τ is that

$$E\left[\varepsilon_{it}H_i\right] = 0.$$

In words, I assume that changes in worker outcomes that are unexplained by my controls are uncorrelated with the type of the worker's destination firm. Economically, this assumption would be satisfied if high and low promotion firms are observationally indistinguishable to workers at the point of job take-up. This assumption would also be satisfied if firms observationally differ, but workers (with the same covariates) draw offers at random from a common distribution and make the same take-up decisions upon receiving an offer. On the other hand, this assumption would be violated if workers who are more likely to experience a positive shock are also more likely to receive or accept an offer from a high promotion firm.

Although my identifying assumption is about the counterfactual *changes* in worker outcomes and not directly testable, a related falsification test suggested by Kahn-Lang and Lang (2020) is whether workers moving to high versus low promotion firms are different in *levels*. Figure 5 plots the difference in baseline covariates between high and low promotion firm movers from the regression

$$X_{i,c-1} = \beta H_i + \theta M_{i,c-1} + \varepsilon_i,$$

where $X_{i,c-1}$ are worker characteristics from the baseline year, H_i is the indicator for whether the worker moved to a high promotion firm, and $M_{i,c-1}$ are any controls from the baseline year. The raw differences in baseline means between the two groups show that movers to high promotion firms have slightly higher baseline earnings, are more likely to be male, and have fewer years of education. However, most of the differences are explained by the state and occupational group of the worker. Any remaining differences in baseline earnings between the two groups are explained by the fact that movers from high promotion firms are more likely to come from high wage firms (as proxied by the estimated AKM firm fixed effect), and remaining differences in education and gender are quantitatively small.

Estimation details

I estimate average employer effects by stacking the five cohorts of workers who make an employerto-employer transition to a high or low promotion firm between 2008 and 2012.¹⁷ Figure 4 shows a timeline of my approach. Since my data end in 2015, I observe at least three years of worker outcomes following the move, but any effects for longer-term outcomes are identified by the earlier cohorts . A recent literature on event studies with staggered treatment timing has cautioned that simply pooling all cohorts and estimating treatment effects by OLS may yield unintuitive and potentially negatively weighted means of the cohort-specific treatment effects (e.g., de Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021; Sun and Abraham, 2021). I avoid this issue by allowing all coefficients in Equation 6 to vary arbitrarily by cohort and then averaging the cohort-specific treatment effects explicitly with uniform weights as $\beta_{\tau} = E \left[\beta_{c\tau}\right]$.¹⁸ This estimation approach also clarifies that the

¹⁷In the cases when the same worker belongs to multiple mover cohorts, I consider the worker's earliest move. However, the results are similar when I include all moves as separate events.

 $^{^{18}}$ This approach is algebraically identical to estimating Equation 6 separately for each cohort. Estimating the coefficients jointly in this approach ensures that the standard errors are correct when clustering across cohorts.

identification of β_{τ} comes solely from aggregating comparisons between workers that are in the same cohort. For more details, see Appendix Section C.2.

I make several sample restrictions on the set of movers to ensure that I capture the effect of interest. First, I restrict the estimation sample to workers who were continuously employed for the three years before the move to facilitate the assessment of pre-trends and to ensure that my sample consists of workers with a high degree of attachment to the formal labor force. I additionally restrict the sample to workers who are between the ages of 25 and 50 at the time that they switch firms. This ensures that my sample includes ages where promotions are most relevant, but worker outcomes are unlikely to be driven by external shocks like schooling and retirement.¹⁹ Finally, I focus on workers for whom promotions would be relevant – production workers working in an occupational group with a directly observable supervisor track in the year before the move. Some employer-to-employer transitions in the data appear to reflect firm organizations or spinoffs, so I use a worker-flows approach to eliminate employer changes that appear to be spurious.²⁰ The remaining transitions reflect worker-level variation and do not require higher clustering under the framework of Abadie et al. (2017), but I cluster standard errors conservatively at the destination firm level to allow for arbitrary correlations in outcomes between workers moving to the same firm and over time.

4.3 Extensions incorporating additional variation

The exogenous job movers assumption that is also sufficient for my identification assumption makes strong restrictions on worker mobility. Although the restrictions are generally consistent with the data, researchers have also pointed out that worker sorting can violate the identifying assumptions but generate similar empirical patterns (for example, see Eeckhout and Kircher, 2011). To address these concerns, I extend the baseline design to incorporate two sources of aggregate variation that relax the two key components of the assumption – that workers separate from their jobs for exogenous reasons, and that conditional on separating, the type of firm the worker moves to is exogenous. In the framework of Abaluck et al. (2021), I test the balance and fallback conditions, respectively. It is worth noting that my approach is not specific to promotions or the Brazilian data, and it may be generally useful as tests for the validity of movers research designs.

To ensure that my effects are not driven by workers' differential selection into moving employers, I follow Gibbons and Katz (1992) and use mass layoffs as an exogenous shock that separates workers from their current employers. I follow the methodology from Schmieder et al. (2020) to identify mass layoff events as firm events where employment dropped by at least 30% year over year and where no more than 20% of separated employees were re-employed at the same firm. Around 127,000 workers from 16,000 firms in my baseline sample were separated as a result of a mass layoff, which is 12% of the baseline sample. Column 2 of Table 5 summarizes the characteristics of the laid-off workers. Relative to the baseline sample, they are slightly older, have fewer years of education, and lower earnings, but they are more likely to move to high promotion firms.

I test for whether selection into moving firms is driving my results by estimating the baseline Equation 6 on the subsample of exogenously separated workers. It is worth emphasizing, however, that I compare laid-off workers who move to high promotion firms to laid-off workers who move

 $^{^{19}}$ Figure A.3 plots the empirical distribution of supervisors and promotions. Approximately 70% of all promotions are between the ages of 25 and 50.

 $^{^{20}}$ Specifically, I follow the mass layoffs literature (e.g., Schmieder et al., 2020) by dropping any origin firms where at least 20% of exiters went to the same destination firm.

to low promotion firms. So, I use mass layoffs as a particularly powerful instrument that shifts all affected workers into moving, but I assume that the effects of the mass-layoff itself are homogeneous and absorbed by my time-varying controls. This is in contrast to the literature on estimating the effects of mass layoffs by comparing laid-off workers to workers at other, non-layoff firms.²¹

To ensure that my effects are not driven by workers differentially sorting into high promotion firms based on idiosyncratic shocks, I consider aggregate regional variation in the types of employers that are hiring workers (which is similar in spirit to Oreopoulos et al., 2012). I construct z_m for each cohort as the jack-knife share of hiring at the worker's municipality that is from high promotion firms rather than low promotion firms, excluding any hires from the worker's destination firm.²² I then instrument H_i with z_m by estimating the following just-identified system of equations using two stage least squares:

$$H_{i} \times I_{t}^{\tau} = \phi_{\tau} \left(z_{m} \times I_{t}^{\tau} \right) + \tilde{\pi} X_{it} + \tilde{\alpha}_{i} + \tilde{\theta}_{t} B_{i,c-1} + \tilde{\gamma}_{sot} + \eta_{it}$$
$$y_{it} = \sum_{\tau} \beta_{\tau} \left(H_{i} \times I_{t}^{\tau} \right) + \pi X_{it} + \alpha_{i} + \theta_{t} B_{i,c-1} + \gamma_{sot} + \varepsilon_{it}.$$
(7)

The second stage equation is identical to the baseline Equation 6, including the use of worker fixed effects and flexible controls for skill prices. Since variation is at the municipality level, I differ from the other specifications by clustering standard errors by the worker's destination municipality.

Figure A.4 shows graphical intuition for my instrument by plotting the average H_i , the probability that a worker moves to a high promotion firm, against binned values of z_m , the jack-knifed municipal hiring share. I control for observable differences in worker characteristics (a quadratic in age interacted with gender along with indicators for the worker's education, race, and state), the firm wage premia of the worker's origin firm, and state-by-occupation fixed effects. z_m has clear predictive power and the conditional expectation function is approximately linear. As an additional check, I estimate the cross-sectional regression

$$H_i = \phi z_m + \tilde{\pi} X_{ic} + \tilde{\theta} B_{i,c-1} + \tilde{\gamma}_{so} + \nu_i \tag{8}$$

separately for each cohort. The minimum F-statistic for the instrument across all five cohorts is 627, and the average F-statistic is 1161.

The key identifying assumption for the instrumental variables approach is:

$$E\left[\varepsilon_{it}z_m\right] = 0.$$

In words, I assume that changes in worker outcomes that are unexplained by my time-varying controls are uncorrelated with the worker's municipality. This municipality-level assumption substantially relaxes the worker-level assumption in Section 4.2 by allowing workers to sort arbitrarily across firms

 $^{^{21}}$ Of course, a concern with this approach is that mass layoffs themselves differentially affect workers. This concern is mitigated in my setting by the fact that my comparison is within the set of mass layoff workers, and also by the fact that any worker who is included in the employer-to-employer mover sample would have been unemployed for less than 12 months following the mass layoff.

²²The workers in my analysis sample contribute to the total number of new hires, which introduces a functional dependence between H_i and the hiring share z_m . Although this problem is minor in my setting since municipalities in the sample are generally large, I exclude any hiring from the worker's destination firm from z_m to avoid the reflection problem that would otherwise arise (Manski, 1993). Technically, the jack-knife procedure introduces some slight variation in z_m by the worker's destination firm, but I slightly abuse notation to make the source of identification clear.

within municipalities, including in anticipation of worker-level changes. However, this assumption also imposes the additional exclusion restriction that z_m is uncorrelated with any other unobserved factors that may shift worker outcomes. This concern is mitigated by the rich state-by-occupationalgroup-by-time fixed effects and testable pre-trends. Moreover, I focus on assessing whether some firms are more likely to generate positive outcomes (promotions) as well as negative outcomes (exits). To the extent to which regions with more hiring by high promotion firms experience different local labor market shocks, these additional shocks are unlikely to increase both worker promotions and formal labor force exits. Nevertheless, potential violations of the exclusion restriction are an important caveat to interpreting the IV results.

5 Direct effects of high promotion firms

My estimates match the model's predictions on the direct effects of employer learning: high promotion firms have persistently positive effects on worker promotions and negative effects on formal labor market attachment. Consequently, these firms have only negligible long-term effects on average worker earnings. I begin by reporting the estimates from the baseline research design described in Section 4.2, which are my preferred estimates since they most directly reflect the effects for the typical worker making a job-to-job transition. To address any remaining concerns that my results may be driven by unobservable differences between workers, I next discuss the results of a series of robustness checks and alternate identification strategies. I show that including rich time-varying trends by additional worker characteristics or removing most of the controls from the baseline specification has little effect on my estimates. I also show that alternative identification strategies that incorporate quasi-experimental shifters of workers' firm choices, as discussed in Section 4.3, produce estimates that are quantitatively and qualitatively similar to my baseline results.

5.1 Baseline results

My first result is that differences in firm promotion rates reflect persistent differences in firms' causal effects on workers. Figure 6 plots the estimates of β_{τ} from estimating Equation 6 on the outcome of whether a worker is working as a supervisor τ years after moving.²³ Workers are 1.19 (s.e. .09) percentage points more likely to be working as a supervisor within the first two years after moving to a high promotion firm rather than a low promotion firm. The effects are large relative to the average promotion rate of 1.68 percentage points for movers to low promotion firms over the same period, as well as the overall annual promotion rate of 0.779 percentage points for production workers. Moreover, the effects are persistent. Workers are still 0.936 (s.e. 0.08) percentage points more likely to be working as a supervisor after initially moving, even though 77% of workers are no longer working at the destination firm by then.²⁴

 $^{^{23}}$ For this outcome, I include all workers who leave the formal labor market, and I assume these workers are not working as a supervisor. The effects could be considered a lower bound if high promotion firm movers who leave the formal labor market are also more likely to work as a supervisor in the informal labor market.

²⁴Note, however, that these effects do *not* necessarily imply that the long term promotion effects are driven by a single group of workers who were immediately promoted. Figure A.5 plots the event study coefficients from the same estimating equation on the outcome of whether a worker was *ever* promoted by τ years after moving. Promotions within a year of moving the firm account only for half of the cumulative number of promotions. So, the relative stability of the effects on promotions at τ masks the fact that workers are entering and exiting the supervisor role at roughly similar rates in the later years.

This result rules out two alternative explanations for the role of high promotion firms. First, promotions are not due to the realization of transient firm shocks. Firms are classified as high or low promotion firms based on their promotion rates for workers between 2004 and 2006, whereas workers in the analysis sample joined the firms between 2008 and 2012. The fact that firms that are active in promotions from the early data also have clear effects on workers who join the firm several years afterward indicates that these effects are driven by persistent differences between firms.²⁵ Second, promotions at high promotion firms cannot be determined by seniority rules alone. My identification strategy relies upon comparing workers around employer transitions, so all workers in the sample would be at the bottom of the seniority ladder. The fact that I detect large effects on promotions immediately after the worker moves firms indicates that firms are willing to consider relatively new workers for more senior positions.

My second result is that high promotion firms are also more likely to lead workers to leave formal employment altogether. Figure 7 plots the baseline event study specification for the outcome of whether the worker is working for *any* firm in the formal employment sector τ years after moving. Workers are 1.93 (s.e. 0.25) percentage points less likely to be found in formal employment more than two years after moving to a high promotion firm rather than a low promotion firm. Although the effects are smaller in relative terms given that 21.4% of workers who move to low promotion firms also leave formal employment over the same period, the differential effect of high promotion firms on exit rates is twice as large as their effect on promotion rates. So, a production worker who joins a high promotion firm is substantially more likely to leave formal employment altogether than to become a supervisor.

The key prediction from Proposition 2 is that workers joining high learning firms are both more likely to be promoted and more likely to become unemployed. The results on employers' effects in this section confirm this prediction. In other words, *high promotion* firms seem to be *high learning* firms. Moreover, the details of the empirical results support the model's assumptions. I assume in the model that the matching of workers to firms is random. Focusing on a sample where this assumption is plausible (based on the evidence in Figure 5) and additionally isolating the effects of the high promotion firms help ensure that the connection between the model and the empirical support is tight. In addition, I model learning as the result of uninformed "trial promotions" that generate information about worker ability. The fact that promotion effects appear immediately after workers move firms and then partially decay over time is consistent with this formulation.

Finally, high promotion firms have only mixed effects on workers' long-term earnings. Figure 8 plots the event study estimates for the worker's average monthly earnings conditional on remaining in formal employment. Workers who move to a high promotion firm earn 1.26 (s.e. 0.51) log points more within the first two years of moving, which is consistent with higher overall wage policies at those firms. However, the short term wage gains dissipate quickly, and those workers earn no more than workers who initially moved to low promotion firms three years after the initial move. Given that earnings are defined only for workers who remain in formal employment, the estimates can be considered an upper bound on the overall effects on earnings, since workers who move to high promotion firms are also more likely to leave formal employment afterward.²⁶ The event study on

 $^{^{25}}$ Moreover, Figure A.6 shows that these effects are not substantially different between the early and later cohorts. 26 Table A.2 provides a lower bound on earnings effects by imputing earnings outcomes for workers who leave the formal labor force under the assumption that all workers who leave are unemployed. To implement this, I assume that counterfactual earnings for workers who leave the formal labor market would have grown at the average rate for

earnings is also likely to be a particularly sharp test of my identifying assumptions. If workers who move to high promotion firms are more likely to have systematically higher productivity growth, then they are likely to have differentially higher earnings growth in the three years before moving; my results show that this is not the case.

5.2 Robustness of direct effects

Several checks support the interpretation that my estimates reflect high promotion firms' causal effects rather than differences in worker sorting or market level shocks. The estimated effects remain quantitatively and qualitatively similar when I use a simplified specification with only worker and year fixed effects or a rich specification with time trends that absorb additional differences between workers at baseline. The estimates are also comparable when I use shifters of worker separations or worker destinations, as discussed in Section 4.3, although these estimates are generally less precise. Finally, the effects do not depend on the details of how I classify firms, but the effects of firms with high promotion rates are distinct from the effects of firms with high wage growth.

Figure 9 compares the pooled estimates of the main outcomes across three different specifications of controls. Compositional differences between high and low promotion firm movers are small in magnitude and become largely indistinguishable from zero conditional on the worker's baseline geography, firm wage premia, and occupation, which motivated controlling for flexible trends at the state-by-occupation level and by baseline firm wage premia. However, dropping these additional controls only slightly decrease the estimated effects on workers' formal labor market attachment and earnings. Meanwhile, adding additional flexible trends for each gender-by-education cell, the worker characteristics for which the difference between movers was statistically significant, as well as flexible trends by baseline worker earnings does not meaningfully change the estimated effects. The stability of estimates across these controls is reassuring that my identification assumption appears valid.

Meanwhile, Figure 10 compares the pooled estimates of the main outcomes across the three sources of variation. The mass layoff estimates are from estimating Equation 6 on the subsample of movers whose baseline firm was experiencing a separation shock that is plausibly exogenous to the worker. The estimates are less precisely estimated than the baseline due to the smaller sample but are otherwise nearly identical. The estimated earnings effects are slightly higher than the baseline estimates, although the two are not statistically distinguishable.²⁷ Meanwhile, the local hiring estimates are from estimating the instrumental variables Equation 7. Variation is at the municipality level, and the standard errors are correspondingly larger, but the estimates remain comparable to my baseline results. The hiring instrument is strong in the full specification – the F-statistic for the joint significance of the excluded instruments in the pooled IV first stage regression is 119, so I calculate confidence intervals based on the standard asymptotically valid critical values (as suggested by Lee et al., 2021).²⁸ The magnitudes of the IV estimates on formal labor market attachment and earnings gains are both larger, and the IV estimate on promotions is again nearly identical to my baseline estimates.²⁹

their baseline occupation-by-state cell, but are discounted by the ratio between the value of non-work time and the market wage (estimated as 0.58 by Mas and Pallais, 2019).

 $^{^{27}}$ Figures A.12, A.13, and A.14 replicate the main event study figures for this subsample. The outcomes are free of pre-trends and the dynamics are similar to that of the main specification.

 $^{^{28}}$ The F-statistic for the earnings outcome is 105. It is slightly different since the sample is conditional on remaining in formal employment.

²⁹Figures A.15, A.16, and A.17 replicate the main event study figures for the IV specification. The effects are noisier

Finally, I find that the results do not depend on the details of how I classify firms based on promotion rates, but the use of actual promotions data is critical. Figures A.9, A.10, and A.11 replicate the event study estimates of the baseline Equation 6 using the continuous measure of the firm's promotion propensity η_j (as defined in Section 4.1 and winsorized at the 5th and 95th percentiles) instead of the discrete high versus low classification. The qualitative patterns are nearly identical when using this continuous measure, and the magnitudes of the estimates are consistent with the overall difference in average promotion rates between high and low promotion firms. Another possible concern is that the classification process includes both external and internal promotions, so the estimated effects are not necessarily representative of firms that focus on developing their workers. I test for this explanation by replicating my baseline empirical strategy on an alternate definition of high and low promotion firms that uses information from internal promotions alone. Figure A.18 compares the pooled estimates from this alternative classification, and the effects are again similar.³⁰

To highlight the importance of direct occupational data, I also benchmark the effects of high promotion firms against the effects of high *wage growth* firms (defined as firms in the top quartile of wage growth for incumbent workers between 2004-2006) in Figure A.18. Unsurprisingly, firms with high wage growth do increase the earnings of their workers, but they have almost no effect on workers' career progressions and they have similarly negative effects on the probability that workers remain in formal employment. This suggests that the effects of high wage growth firms may primarily be driven by other channels of wage determination like bargaining, on-the-job search, or seniority pay.

6 Additional evidence for the employer learning mechanism

The analysis in Section 5 shows that the effects of high promotion firms on worker promotions, labor force exit, and earnings are robust to alternate specifications and unlikely to be driven by worker sorting. Furthermore, given the setup of the empirical strategy, the effects are unlikely to be explained by transient shocks or seniority rules at high promotion firms. In this section, I use a variety of additional empirical exercises to show that the effects of high promotion firms are consistent with employer learning and inconsistent with several other alternative explanations.

6.1 Promotion effects reflect positive worker outcomes

I empirically examine whether differences in promotions reflect differences in real outcomes by comparing the promotion wage premium for workers who move to high promotion firms versus low promotion firms. To do so, I extend Equation 6 and estimate

$$y_{it} = \beta_{\tau}^{HP} \left(H_i \times P_{i,c+1} \times I_t^{\tau} \right) + \beta_{\tau}^H \left(H_i \times (1 - P_{i,c+1}) \times I_t^{\tau} \right) + \beta_{\tau}^P \left(P_{i,c+1} \times I_t^{\tau} \right) + \pi X_{it} + \alpha_i + \theta_t B_{i,c-1} + \gamma_{sot} + \varepsilon_{it}$$

$$(9)$$

⁽especially for the promotion and exit outcomes), but the dynamics again remain similar. There is some evidence that workers in regions with more high promotion employers are more likely to be working as a supervisor in the years before the move, but this difference is less than a third of the immediate promotions effects. Importantly, those workers also do not appear to be on differential earnings trajectories in the years before the move.

 $^{^{30}}$ The similarity of these two effects is unsurprising since most of the promotions that determined a firm's promotion propensity were internal promotions. For more details on classification, see Appendix C.1.

where $P_{i,c+1} = 1$ if worker *i* is working as a supervisor one year after moving firms and $P_{i,c+1} = 0$ otherwise. These estimates are descriptive rather than causal since promotions are outcomes from moving to high promotion firms. But the exercise is useful for testing for the patterns of selection that would result if high promotion firms are simply more likely to classify workers as supervisors.

Workers who are promoted following a move to a high promotion firm experience persistent increases in earnings that are at least as large as the promotion wage-premia at low promotion firms. Figure 11 plots the estimates of $(\beta_{\tau}^{HP}, \beta_{\tau}^{H}, \beta_{\tau}^{P})$, the relative wage changes for promoted high promotion firm movers, unpromoted high promotion firm movers, and promoted low promotion firm movers, respectively. The omitted reference group is the unpromoted low promotion firm movers. Promotions are meaningful for these workers – high promotion firm movers who are promoted within a year of moving receive wage increases of 15.6 (s.e. 0.97) log points, while corresponding low promotion firm movers receive wage increases of 11.9 (s.e. 0.66) log points. These effects decay over time, but remain economically and statistically significant even seven years after the move. Meanwhile, high promotion firm movers who were not promoted within a year of moving initially earn slightly more than corresponding low promotion firm movers, but their earnings are nearly identical more than two years after the move.

These results are inconsistent with the alternative explanation that the promotion effects reflect only surface-level firm differences in the use or reporting of job titles.³¹ Under this explanation, both promoted and unpromoted workers at high promotion firms would be more negatively selected than the same groups at low promotion firms. As a concrete example, suppose the distribution of workers' wage growth Δw is $F(\Delta w)$ for both high and low promotion firms, and a firm f labels a worker as a supervisor when Δw exceeds some cutoff c_f . A high promotion firm H is a firm with a lower cutoff for promoting workers (i.e., $c_H < c_L$), so it follows that

$$E [\Delta w | \Delta w > c_H] < E [\Delta w | \Delta w > c_L]$$
$$E [\Delta w | \Delta w \le c_H] < E [\Delta w | \Delta w \le c_L] .$$

My results from Figure 11 are the opposite, which suggests that differences in promotions reflect real differences in job assignments (and corresponding wage structures) rather than differences in the willingness to report similar jobs as supervisory jobs.

6.2 Exits from formal labor force are involuntary

The main results in Section 4 show robust evidence that high promotion firms increase the likelihood that workers exit from the formal labor force. Although employer-to-unemployment transitions are a common proxy for involuntary turnover (for example, in Moscarini and Postel-Vinay, 2018; Sorkin, 2018), the large informal sector in Brazil may complicate their interpretation in my setting. I present two additional pieces of evidence that bolster the interpretation that increased exits reflect the higher use of separations by high promotion firms rather than alternate explanations like different outside options or voluntary quits.

The first result is that the higher formal labor force exit effects are consistent with higher firm

 $^{^{31}}$ Although the use of administrative data mitigates some of these concerns, worker occupations are self reported by firms, and there may be additional incentives to classify workers differently due to collective bargaining agreements or the use of job titles as status (Baron and Bielby, 1986; Lagos, 2019).

separations at high promotion firms. Table A.2 presents the pooled effects of high promotion firms on the probability that the worker is still working at the destination firm. Workers are 4.47 (s.e. 0.82) percentage points more likely to leave their destination firm more than two years after moving to a high promotion firm instead of a low promotion firm, which is more than double the effect on formal labor force exit. However, firm exit is a less direct measure of negative employment outcomes than formal labor force exit, so I use it as a supporting fact rather than the main outcome. Nevertheless, it is reassuring that the effects on formal sector exit can be rationalized by effects of separations from the destination firm.

In addition, I directly test for whether high promotion firms affect worker-initiated turnover and find a precise zero effect. A feature of RAIS is that it records the reason why an employment contract is terminated, including whether the termination was initiated by the firm or by the worker. This distinction matters in Brazil because the employer pays higher separation penalties for employer terminations without cause (including layoffs). I classify employer terminations with or without cause (i.e., firings or layoffs) as firm-initiated separations, and I classify voluntary quits as worker-initiated separations. I observe reasons for separation preceding around half of spells outside of formal employment, of which 77.6% are firm-initiated and 14.5% are worker-initiated.³² Figure 12 reports the estimated effect of moving to a high promotion firm on the likelihood of each type of exit. Firm-initiated exits account for virtually all of these additional separations, whereas high promotion firms' effects on worker-initiated exits are precisely zero.

6.3 Promotion and exit effects are stronger for high potential workers

To test whether promotions mediate both positive and negative job outcomes, I compare the effects of high promotion firms for workers who are ex-ante more likely versus less likely to be promoted. Abadie et al. (2018) cautions that even using prior characteristics to predict subsequent outcomes can introduce bias due to overfitting, so I split my analysis sample into a 25% hold-out sample and a 75% estimation sample. I first estimate the equation

$$Promoted_{i,c+1} = \theta Z_{i,c-1} + \varepsilon_i \tag{10}$$

on the 25% hold-out sample of movers. I then rank workers in the 75% estimation sample by their predicted promotion potential $(\hat{\theta}Z_{i,c-1})$, and I separately estimate the baseline event study Equation 6 for the top and bottom tercile of workers by promotion potential. Table A.3 reports summary statistics for the high and low potential worker subsamples.³³ As expected, high potential workers have more years of education and higher baseline earnings than low potential workers. They are also more likely to work as supervisors and less likely to exit.

Figure 13 compares the main effects of high promotion firms on worker promotions, formal labor force exit, and average earnings for the baseline, high promotion potential, and low promotion potential samples. The fact that the effects on worker promotions are strongest for the high potential workers is reassuring but unsurprising. On the other hand, the fact that the effects on formal labor force exit are also larger for these high potential workers contradicts the alternative explanation that

³²The remaining reasons for separations are primarily contract expirations, transfers, and retirements.

 $^{^{33}}$ I define worker potential terciles prior to making the additional sample restrictions discussed in Section 4.2 to ensure that the definitions would be stable across alternate sample restrictions. This introduces some differences between the size the two subsamples but does not affect the interpretation of the effects.

high promotion firms simply have more volatile labor demand. For example, firms with "dual labor markets" may offer some workers both job security and promotion opportunities and other workers precarious jobs that are eliminated during demand downturns. My results instead suggest that promotions and exits are related outcomes that are two sides of the same coin. These results lend further support to employer learning as the key mechanism that rationalizes the employer effects.³⁴

6.4 Survey of firm labor practices is consistent with estimated effects

Finally, I validate my worker-level evidence using structured interview data from an internationally comparable survey of manufacturing plants' human capital management practices.³⁵ I view these data as supporting evidence that my results about the worker-level effects on promotions and exits are consistent with managers' actions. Furthermore, the international comparability of the survey allows me to assess the extent to which the patterns are specific to Brazil

Two questions from the World Management Survey connect to the worker-level outcomes that I study. Table A.4 describes each question from the questionnaire and its corresponding scoring criteria. The question on promotions from the questionnaire is primarily about the *reason* for promotion rather than the *frequency* of promotions. However, in my setting, a firm that primarily promotes workers based on tenure should have negligible estimated effects on promotions since my sample focuses on workers who move firms and consequently are at the bottom of the seniority ladder. Meanwhile, the question on firings directly connects to my results on formal labor force exits given the discussion in Section 6.2.

Figure 14 plots the average firings score by firms' promotions score for three groups of countries. First, I find that firms that are more likely to develop and promote high performing workers are also more likely to fire low performing workers. This is consistent with my baseline empirical results, as well as the general finding in the management and organization literatures that optimal personnel policies tend to be a suite of complementary practices (Milgrom and Roberts, 1992; Ichniowski et al., 1997; Bloom et al., 2019; Benson et al., 2019; Cornwell et al., 2021).³⁶ Second, the correlations between firms' promotions practices and firing practices are similar between Brazil, OECD countries, and non-OECD countries.³⁷ These three groups differ substantially in their labor market institutions, industrial compositions, and the allocation of production factors. The similarity of the relationship across all three settings further supports the interpretation that my results are driven by a general economic mechanism rather than any features that are specific to Brazil.

 $^{^{34}}$ In the model, the effects on high potential workers can be rationalized by workers for whom trial promotions are more relevant (so Q is higher). On the other hand, it would be difficult to explain this pattern without learning (e.g., differences in the types of workers at different firms) since high potential workers are generally *less* likely to leave formal employment.

³⁵There are general caveats to interpreting survey responses since managers' reported promotions practices may reflect over-optimism or worker sorting rather than the firm's true causal effects. Some of these concerns are mitigated by the survey I use, which comes from a series of structured interviews between MBA students and plant managers that are designed to minimize incentives to misreport firm practices. For more details about the survey, see Bloom et al. (2014).

³⁶The correlation between these two practices is also notable since analyses of internal labor markets usually argue that investment is more likely when turnover is low (for example, see Doeringer and Piore, 1971; Prendergast, 1993).

³⁷The set of OECD countries in the sample are Australia, Canada, Chile, Columbia, France, Germany, Great Britain, Greece, Italy, Japan, New Zealand, Poland, Portugal, Ireland, Spain, Sweden, United Kingdom, and the United States. The set of non-OECD countries in the sample are Argentina, China, Ethiopia, Ghana, India, Kenya, Mexico, Mozambique, Myanmar, Nicaragua, Nigeria, Singapore, Tanzania, Turkey, Vietnam, and Zambia.

7 Equilibrium Effects of Employer Learning

In addition to direct effects on the labor market outcomes of their employed workers, high learning firms also have equilibrium effects on the wage structure for all workers. To fully explore these implications, I endogenize the employment share of high learning firms and the secondary market contact rate from the partial equilibrium model in Section 2. I follow Lise and Robin (2017) in assuming that firms set the number of vacancies based on expected profits, and that the total cost of vacancies is convex. If the vacancy supply function is sufficiently convex (e.g., if it is log-convex), then high learning employers post a proportionately larger share of vacancies in more productive areas. A higher share of high learning employers increases the adverse selection of job movers, which decreases market wages for promoted and especially unpromoted workers. As a result, occupational wage inequality increases. I show that these equilibrium predictions also hold robustly in the data – municipalities with more high learning employers have higher wage differentials between promoted and unpromoted workers even after controlling for a wide range of worker-, job-, and municipality-level characteristics.

7.1 Endogenizing vacancy creation

In addition to the baseline setup, suppose that locations differ in a multiplicative productivity term, ψ , so output in region r for a worker of type θ_i in occupation o is

$$f_{roi} = \psi_r f_o\left(\theta_i\right)$$

There is an equal mass of atomistic firms, each choosing a density of vacancies v at a total cost of c(v) at the start of the period. The matching technology is constant returns to scale and does not distinguish between the type of firm offering the vacancy or the type of worker. Initial job matches are formed through matching workers with vacancies. Similarly, contact with the secondary market is established by matching initially employed workers with initially unfilled vacancies, so the number of workers who will make contact with the secondary market can be written as

$$m^S = M(m^I, v - m^I)$$

where $m^{I} = M(l, v)$ is the mass of initially employed workers, l is the total mass of workers, v is the total mass of vacancies, and M is the matching function. I assume that, conditional on arriving at the secondary market, firms still compete for workers as in Section 2.2.

Implicit in this setup are two simplifying assumptions to ensure the model remains tractable. First, I rule out any subsequently vacated positions (e.g., from a firm that fired its initially matched worker) from also joining the secondary market so that employment shares always correspond to vacancy shares. Second, I rule out any initially unmatched workers joining the secondary market so the initial level of vacancies will not affect the equilibrium degree of adverse selection.

The Perfect Bayesian equilibrium in this extended model is a set of worker and firm strategies such that each firm sets the number of vacancies that maximizes expected profits given the vacancy setting decisions of other firms, and such that the conditions in the partial equilibrium model in Section 2.2 are also satisfied. Proposition 3 compares the different outcomes that arise in equilibrium between regions that differ only in general productivity. **Proposition 3.** Suppose (i) vacancy creation as a function of expected profits is log-convex and (ii) the elasticity of re-employment with respect to the number of high learning employer vacancies exceeds the elasticity of expected output for unknown secondary market workers. Then, in more productive regions: (i) high learning employers post a greater share of vacancies, (ii) offered wages for incumbent workers are higher (iii) wage differentials between promoted and unpromoted workers are larger.

The full proof for Proposition 3 is in Appendix B and relies on two assumptions. The first assumption ensures that vacancy creation is sufficiently responsive so that increases in the difference in expected profits between high and low learning employers will increase the share of total vacancies created by high learning employers. The second assumption ensures that the net effect of an additional vacancy from high learning firms is positive for workers' expected wages.

The need for the second assumption highlights the negative equilibrium effects of high learning employers on wages. From Equation 4, offered wages at incumbent firms match expected offers in the secondary market. Holding the share of high learning employers equal, increasing the total number of vacancies would raise expected outside options by increasing the likelihood that workers encounter the secondary market. On the other hand, holding the total number of vacancies equal, increasing the share of high learning employers would exacerbate the adverse selection of workers entering the secondary market. Secondary market wages for unknown workers would correspondingly fall, which lowers expected outside options for promoted and especially unpromoted workers. The second assumption in Proposition 3 ensures that the positive wage effects of the vacancy dominate the negative wage effects of high learning employers for unpromoted workers. However, wage inequality would increase in either case since promoted workers are more insulated from the beliefs of the secondary market due to the informativeness of their job assignments.³⁸

7.2 Testing for equilibrium effects on occupational wage structure

Proposition 3 also yields testable implications about the relationship between a region's high learning employer share and its market-level wages. As an empirical analog, I define

$$\bar{H}_m = \frac{\sum_{Q_f = Q_H} L_f}{\sum_{Q_f = Q_H} L_f + \sum_{Q_f = Q_L} L_f}$$

as the (employment-weighted) share of high promotion firms in a municipality, and I estimate the regression

$$y_{imt} = \beta_1 \bar{H}_m + \beta_2 \left[\bar{H}_m \times P_{it} \right] + \theta_1 X_{it} + \theta_2 \left(Z_m, Z_m \times P_{it} \right) + \sigma_{op} + \psi_f + \gamma_{sot} + \epsilon_{it}$$
(11)

on a longitudinal 5% sample of all promotion-track workers between the ages of 25 and 50 in Brazil. P_{it} is an indicator for whether the worker is working as a supervisor, X_{it} controls for worker characteristics (a quadratic in age interacted with gender along with indicators for the worker's education, race, and state), and Z_m controls for region size. Finally, σ_{op} is a full set of occupational-group-by-

³⁸Although I motivate differences in high learning employer shares as arising from vacancy creation in response to productivity differences, neither of those components are crucial for this mechanism. The equilibrium channel where high learning employers suppress wages and increase wage inequality would exist whenever wage determination follows the structure from the partial equilibrium model.

supervisor fixed effects, ψ_f are firm fixed effects, and γ_{sot} are state-by-occupational-group trends.

Table 6 reports the estimates of (β_1, β_2) , the correlation between regional wages and the high promotion employment share, as I progressively add the controls in Equation 11. Both β_1 and β_2 are positive and substantial in the basic specification that controls only for flexible trends in wages at the state-by-occupation level. Adding additional controls generally has little effect on the estimates despite their substantial explanatory power on the variance in wages, which suggests that the estimates of (β_1, β_2) are not driven by observable differences between the characteristics of jobs, workers, or municipalities. The one exception is that the estimate of β_1 , the overall regional wage premium, becomes small and insignificant upon the inclusion of firm fixed effects, but the regional supervisor wage premia remain substantial. Under my preferred specification that controls flexibly for worker characteristics and different wage structures within occupations (column 3), moving from the 10th to the 90th percentile in the population-weighted municipal high promotion firm share (from 17.3% to 50.8%) increases regional wages for production workers by 4.69 log points and the region wage premium for supervisors by an additional 6.57 log points, which is 39% of the average supervisor wage premium from Table 2.

8 Structural quantification

The model in Section 2 and its extension in Section 7 are both highly stylized to clarify the mechanism of employer learning and characterize its equilibrium implications. However, the model's basic components governing learning and job assignment also have empirical content, and they can be mapped to the data under minimal assumptions about wage setting or the labor market parameters. In this section, I show how those key parameters can be identified using the employer effects estimates under internally consistent model restrictions. The resulting estimates allow me to quantify the degree of employer learning by high and low promotion firms as well as the implied degree of skill misallocation and adverse selection in my setting.

8.1 Model identification and estimation

The partial equilibrium model in Section 2 can be separated into two parts. The first half of the model characterizes the incumbent employer's learning and job assignment decision while taking market wages as given. The second half of the model specifies the equilibrium wage from the secondary market. This separability between the two parts is crucial for transparently mapping my estimates to the learning and job assignment problem while remaining agnostic about the exact form of wage setting. The key condition is that employers must find the efficient job assignments to be incentive compatible. In other words, employers must find it optimal to promote workers of known high ability, fire workers of known low ability, and retain workers of unknown ability. This incentive compatibility condition would be satisfied under a variety of wage setting institutions that would be relevant in Brazil, including rent sharing, wage bargaining, wage floors, or firm-wide wage schedules. As a result, predictions from the learning and job assignment problems are particularly likely to be robust to departures from the stylized setting of the model, and I focus on quantifying their relevant parameters.

To further bolster the transparency and robustness of the quantification exercise, I use a methods of moments approach for estimation. The approach allows me to be explicit about the exact moments from the data that identify the parameters and the relationship between the two (Andrews et al., 2020). I can also choose the moments to reflect the model's assumptions about timing and heterogeneity (similar to the approach from Lamadon et al., 2019). Table 7 summarizes the parameters of interest and the moments used to identify the parameters.

I first calibrate two moments "outside" of the model. I use the probability that a promoted worker remains a supervisor upon moving employers to calibrate the asymmetric information parameter κ . I also use the average job destruction rate estimated in Dix-Carneiro et al. (2021), aggregated to the average length of the outcome period in the sample, to calibrate the exogenous separation rate δ . Since there is greater uncertainty in these moments, I systematically vary the calibrated values to ensure that my conclusions are not sensitive to the choice of these values.

I jointly identify the remaining parameters governing job assignment and turnover $(\bar{Q}_H, \bar{Q}_L, \alpha, g)$ by matching the sample means and treatment effects from the movers analysis using classical minimum distance. The key model restriction that allows me to identify the model is that I assume movers to high versus low promotion firms differ only in the rate of learning at their destination firms. This restriction ensures that the model analogs for the means and treatment effects estimates for promotions and formal labor force exit are determined by the four parameters alone.³⁹ In theory, I can calibrate the model using observational data from all workers in my sample. However, the comparability of baseline characteristics and the absence of systematic pre-trends in earnings between high promotion firm and low promotion firm movers (as discussed in Sections 4.2 and 5) suggest that the model restriction is most plausible in the movers analysis. I use the sample means and treatment effects from at least two years after workers initially move firms to allow enough time for learning and turnover to take place.

It is straightforward to implement the quantification approach since the model-implied analogs of the moments can be computed in closed form. The estimate $\hat{\theta} = (\bar{Q}_H, \bar{Q}_L, \alpha, g)$ is defined as

$$\hat{\theta} = \arg\min_{\theta} \left(\hat{\pi} - h\left(\theta\right) \right)' W\left(\hat{\pi} - h\left(\theta\right) \right), \tag{12}$$

where $\hat{\pi}$ is the vector of empirical moments, $h(\theta)$ is the vector of corresponding model outcomes (given θ), and W is a positive definite weighting matrix. Since $\hat{\pi}$ comes directly from sample calculations, I calculate its full cluster-robust variance-covariance matrix $\hat{\Omega}$ by stacking the estimation equations and clustering the stacked equation by the worker's destination firm. I use the inverse of the variance-covariance matrix as my weighting matrix (so $W = \hat{\Omega}^{-1}$), although I can generally match the empirical moments exactly so the choice of the weighting matrix affects only the efficiency of my estimator. Asymptotically valid standard errors for $\hat{\theta}$ also follow from $\hat{\Omega}$ using the sandwich estimator from Newey and McFadden (1994).⁴⁰ For more details, see Appendix C.3.

³⁹As an example, suppose that both sample means and treatment effects on promotions are large, but sample means and treatment effects on labor force exit are small. The moments on promotions would imply that the relative difference in learning between high and low employers is large, overall learning is sufficiently high, and the share of high ability workers in the population is large. The moments on employment would then distinguish between a high overall rate of learning versus a large share of high ability workers and pinpoint the implied re-employment rate.

 $^{^{40}}$ This approach does ignore the uncertainty due to the calibrated parameters, as discussed by Cocci and Plagborg-Møller (2019). I assess the sensitivity of θ to calibrated values by explicitly assuming alternative calibrations to account for model uncertainty in addition to the (known) sampling uncertainty of the sample means.

8.2 Parameter estimates

Table 8 summarizes the parameter estimates under my baseline calibration, which have intuitive interpretations given the simple structure of the model. I estimate from α that 21.4% of workers in the sample are of high ability. Given that the average share of workers who are in supervisory roles at least two years after initially moving firms is 2.8%, my estimate implies that a high ability worker has an 86.9% chance of being mismatched to a production job or outside formal employment.⁴¹ Part of the reason for the misallocation is because turnover is reasonably high, and information about worker quality is likely to be lost when workers change firms. Under the estimated parameters, a worker who is promoted by their incumbent employer has only a 75.3% chance of remaining in the supervisor position after potential separation shocks and secondary market matching are realized.

But the primary reason for the high rate of mismatch is that the estimated rate of learning for even high promotion employers is 20.8%, so most workers' abilities remain unknown to employers. Although learning is generally low in this setting, it is still substantial compared to the overall rate of worker turnover. As a result, I estimate that adverse selection can be a meaningful source of ex-post market power – workers who change employers are only 70.8% as likely to be of high ability as the overall population of workers in this sample.

Finally, the structural model also gives me a framework to quantitatively assess the overall degree to which high and low promotion firms differ. I estimate that high promotion firms, on average, are 5.8 percentage points more likely than low promotion firms to learn about the ability of their workers. This difference is larger than the treatment effects on either promotions or formal labor force exit, as well as the naive sum of the two effects. Intuitively, the estimated long-run effects on promotions and exit are attenuated by worker turnover and re-employment. The structural model provides a principled approach to adjust for this attenuation using the labor market parameters, which allows me to map differences in long-run outcomes to differences in firm behavior.

Finally, Figure A.19 shows how the main conclusions change when I vary the two calibrated parameters by 50% to 150% of the assumed baseline values. The assumed value for the asymmetric information parameter κ has a quantitatively small effect on my estimates. On the other hand, the assumed rate of exogenous separations δ is important for inferring whether a job mover is a low ability worker who was previously fired or a high ability worker who exogenously separated. Moreover, at 50% of the baseline assumed value for δ , my model can no longer exactly fit the sample means and treatment effect estimates, so those estimates should be interpreted with caution. For all other assumed values of δ , my estimates for skill misallocation and adverse selection remain similar. Since the baseline assumed value of δ is likely a lower bound, the robustness of my conclusions to higher values of δ is particularly reassuring.⁴²

 $^{^{41}}$ A stronger interpretation of the results would be that under perfect information, 21% of the sample would be working as supervisors. However, I do not model the full organizational structure of the firm, and diminishing returns would imply that the actual counterfactual share of supervisory workers would be lower. I view these calculations as more indicative of the degree of imperfect information that a high ability worker faces in the current market.

⁴²The external estimate for δ is the estimated *firm* destruction rate from Dix-Carneiro et al. (2021), whereas the parameter in the model is the exogenous *worker* separation rate. If workers have a positive probability of exogenously separating without the job itself disappearing (e.g., if a worker leaves due to geographic reasons or idiosyncratic preference shocks), then the firm destruction rate would be an underestimate of the exogenous separation rate.

9 Conclusion

I show that employers do influence workers' subsequent careers. Production workers who join high promotion firms are more likely to eventually work as supervisors, but they are also more likely to leave formal employment altogether, and they do not earn more on average. I argue that these results are most consistent with the explanation that employers' information about worker ability at the time of hiring is imperfect, and that employers vary systematically in the degree to which they learn about their workers' abilities. Both treatment effects estimates and a structural quantification show that employer learning is key to rationalizing the effects of employers on worker promotion and exit.

More generally, this paper highlights the role of information, particularly asymmetric information, in determining worker outcomes. Firms acting on information revealed after hiring introduce workerlevel risk in employment that is separate from firm-level or market-level risks. This additional risk has implications for the design of employment protections and unemployment insurance. Adverse selection in the pool of job movers also suppresses workers' outside options, implying that asymmetric information can be a substantial source of ex-post market power for employers. Finally, my model provides an example where the improvement in information about workers exacerbates labor market frictions when that information is privately held. This highlights that the overall effects of more sophisticated employment practices may differ from the direct effects.

The current analysis also points to several fruitful avenues for further research. Adding information on firm accounts can help quantify the rents generated by employer learning as well as the other implications of organizational design. Similarly, adding survey evidence on worker outcomes in the informal sector would help inform the overall welfare trade-offs of high promotion firms. In addition, unions are prominent institutions in Brazil, and it would be useful to assess the degree to which collective bargaining contracts mediate firm practices. Finally, there may be gender and racial differences in the overall effects of firms in my setting; these questions are outside of the scope of this paper, but they are important for understanding who is benefiting from good opportunities and who is getting shaken off the career ladder.

References

- Abadie, Alberto, Susan Athey, Guido Imbens, and Jeffrey Wooldridge (2017) "When Should You Adjust Standard Errors for Clustering?" arXiv:1710.02926 [econ, math, stat].
- Abadie, Alberto, Matthew M. Chingos, and Martin R. West (2018) "Endogenous Stratification in Randomized Experiments," The Review of Economics and Statistics, 100 (4), 567–580.
- Abaluck, Jason, Mauricio Caceres Bravo, Peter Hull, and Amanda Starc (2021) "Mortality Effects and Choice Across Private Health Insurance Plans," *The Quarterly Journal of Economics*, 136 (3), 1557–1610.
- Abowd, John M., Francis Kramarz, and David N. Margolis (1999) "High Wage Workers and High Wage Firms," *Econometrica*, 67 (2), 251–333.
- Acemoglu, Daron and Jorn-Steffen Pischke (1998) "Why Do Firms Train? Theory and Evidence," The Quarterly Journal of Economics, 113 (1), 79–119.

- Addario, Sabrina L. Di, Patrick M. Kline, Raffaele Saggio, and Mikkel Sølvsten (2021) "It Ain't Where You're From, It's Where You're At: Hiring Origins, Firm Heterogeneity, and Wages.," Working Paper 28917, National Bureau of Economic Research.
- Altonji, Joseph G. and Charles R. Pierret (2001) "Employer Learning and Statistical Discrimination," The Quarterly Journal of Economics, 116 (1), 313–350.
- Alvarez, Jorge, Felipe Benguria, Niklas Engbom, and Christian Moser (2018) "Firms and the Decline in Earnings Inequality in Brazil," American Economic Journal: Macroeconomics, 10 (1), 149–189.
- Andrews, Isaiah, Matthew Gentzkow, and Jesse M. Shapiro (2017) "Measuring the sensitivity of parameter estimates to estimation moments," *The Quarterly Journal of Economics*, 132 (4), 1553– 1592.
- (2020) "Transparency in Structural Research," Journal of Business & Economic Statistics, 38 (4), 711–722.
- Arellano-Bover, Jaime (2020) "Career Consequences of Firm Heterogeneity for Young Workers: First Job and Firm Size," 90.
- Arellano-Bover, Jaime and Fernando Saltiel (2020) "Differences in On-the-Job Learning Across Firms," 65.
- Autor, David H. (2001) "Why Do Temporary Help Firms Provide Free General Skills Training?" The Quarterly Journal of Economics, 116 (4), 1409–1448.
- Bagger, Jesper, François Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin (2014) "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics," *American Economic Review*, 104 (6), 1551–1596.
- Baker, George, Michael Gibbs, and Bengt Holmstrom (1994) "The Internal Economics of the Firm: Evidence from Personnel Data," *The Quarterly Journal of Economics*, 109 (4), 881–919.
- Baron, James N. and William T. Bielby (1986) "The Proliferation of Job Titles in Organizations," Administrative Science Quarterly, 31 (4), 561–586.
- Becker, Gary S. (1964) Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, First Edition: NBER.
- Benson, Alan, Danielle Li, and Kelly Shue (2019) "Promotions and the Peter Principle," The Quarterly Journal of Economics, 134 (4), 2085–2134.
- Blau, Francine D. and Jed Devaro (2007) "New Evidence on Gender Differences in Promotion Rates: An Empirical Analysis of a Sample of New Hires," *Industrial Relations: A Journal of Economy* and Society, 46 (3), 511–550.
- Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen (2019) "What Drives Differences in Management Practices?" American Economic Review, 109 (5), 1648–1683.
- Bloom, Nicholas, Renata Lemos, Raffaella Sadun, Daniela Scur, and John Van Reenen (2014) "The New Empirical Economics of Management," Technical Report w20102, National Bureau of Economic Research, Cambridge, MA.
- Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa (2019) "A Distributional Framework for Matched Employee Data," *Econometrica*, 87 (3), 699–739.
- Burdett, Ken and Melvyn Coles (2003) "Equilibrium Wage-Tenure Contracts," *Econometrica*, 71 (5), 1377–1404.
- Caldwell, Sydnee and Nikolaj Harmon (2019) "Evidence from Coworker Networks," 105.

- Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline (2018) "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, 36 (S1), S13–S70.
- Card, David, Ana Rute Cardoso, and Patrick Kline (2016) "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women *," The Quarterly Journal of Economics, 131 (2), 633–686.
- Card, David, Jörg Heining, and Patrick Kline (2013) "Workplace Heterogeneity and the Rise of West German Wage Inequality," The Quarterly Journal of Economics, 128 (3), 967–1015.
- de Chaisemartin, Clément and Xavier D'Haultfœuille (2020) "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," American Economic Review, 110 (9), 2964–2996.
- Cocci, Matthew D and Mikkel Plagborg-Møller (2019) "Standard Errors for Calibrated Parameters," 26.
- Cornwell, Christopher, Ian M. Schmutte, and Daniela Scur (2021) "Building a Productive Workforce: The Role of Structured Management Practices," *Management Science*.
- Cullen, Zoë and Ricardo Perez-Truglia (2021) "The Old Boys' Club: Schmoozing and the Gender Gap," 50.
- Dix-Carneiro, Rafael, Pinelopi Koujianou Goldberg, Costas Meghir, and Gabriel Ulyssea (2021) "Trade and Informality in the Presence of Labor Market Frictions and Regulations," 71.
- Dix-Carneiro, Rafael and Brian K. Kovak (2017) "Trade Liberalization and Regional Dynamics," American Economic Review, 107 (10), 2908–2946.
- Doeringer, Peter B. and Michael J. Piore (1971) Internal labor markets and manpower analysis: Lexington, Mass., Heath [1971].
- Eeckhout, Jan and Philipp Kircher (2011) "Identifying Sorting-In Theory," The Review of Economic Studies, 78 (3), 872–906.
- Farber, Henry S. and Robert Gibbons (1996) "Learning and Wage Dynamics," The Quarterly Journal of Economics, 111 (4), 1007–1047.
- Friedrich, Benjamin (2020) "Internal Labor Markets and the Competition for Managerial Talent," 84.
- Gerard, François, Lorenzo Lagos, Edson Severnini, and David Card (2018) "Assortative Matching or Exclusionary Hiring? The Impact of Firm Policies on Racial Wage Differences in Brazil," 94.
- Gibbons, Robert and Lawrence Katz (1992) "Does Unmeasured Ability Explain Inter-Industry Wage Differentials," The Review of Economic Studies, 59 (3), 515.
- Gibbons, Robert and Michael Waldman (1999) "A Theory of Wage and Promotion Dynamics inside Firms," The Quarterly Journal of Economics, 114 (4), 1321–1358.
- Golan, Limor (2005) "Counteroffers and Efficiency in Labor Markets with Asymmetric Information," Journal of Labor Economics, 23 (2), 373–393.
- Gonzaga, Gustavo, William F. Maloney, and Alejandra Mizala (2003) "Labor Turnover and Labor Legislation in Brazil," *Economía*, 4 (1), 165–222.
- Goodman-Bacon, Andrew (2021) "Difference-in-differences with variation in treatment timing," Journal of Econometrics, 225 (2), 254–277.
- Greenwald, Bruce C. (1986) "Adverse Selection in the Labour Market," The Review of Economic Studies, 53 (3), 325–347.

- Gregory, Victoria (2019) "Firms as Learning Environments: Implications for Earnings Dynamics and Job Search," 57.
- Haanwinckel, Daniel and Rodrigo R Soares (2020) "Workforce Composition, Productivity, and Labor Regulations in a Compensating Differentials Theory of Informality," 93.
- Herkenhoff, Kyle, Jeremy Lise, Guido Menzio, and Gordon M. Phillips (2018) "Production and Learning in Teams," Working Paper 25179, National Bureau of Economic Research.
- Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi (1997) "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines," American Economic Review, 87 (3), 291–313.
- Jarosch, Gregor (2015) "Searching for Job Security and the Consequences of Job Loss," 56.
- Jarosch, Gregor, Ezra Oberfield, and Esteban Rossi-Hansberg (2019) "Learning from Coworkers,"Technical Report w25418, National Bureau of Economic Research, Cambridge, MA.
- Jovanovic, Boyan (1979) "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 87 (5), 972–990.
- Jovanovic, Boyan and Yaw Nyarko (1997) "Stepping-stone mobility," Carnegie-Rochester Conference Series on Public Policy, 46, 289–325.
- Kahn, L. B. and F. Lange (2014) "Employer Learning, Productivity, and the Earnings Distribution: Evidence from Performance Measures," *The Review of Economic Studies*, 81 (4), 1575–1613.
- Kahn-Lang, Ariella and Kevin Lang (2020) "The Promise and Pitfalls of Differences-in-Differences: Reflections on 16 and Pregnant and Other Applications," *Journal of Business & Economic Statis*tics, 38 (3), 613–620.
- Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar (2019) "Who Profits from Patents? Rent-Sharing at Innovative Firms," The Quarterly Journal of Economics, 134 (3), 1343–1404.
- Lagos, Lorenzo (2019) "Labor Market Institutions and the Composition of Firm Compensation: Evidence from Brazilian Collective Bargaining."
- Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler (2019) "Imperfect Competition, Compensating Differentials and Rent Sharing in the U.S. Labor Market," SSRN Electronic Journal.
- Lange, Fabian (2007) "The Speed of Employer Learning," Journal of Labor Economics, 25 (1), 1–35.
- Lazear, Edward P. (1979) "Why Is There Mandatory Retirement?" Journal of Political Economy, 87 (6), 1261–1284.
- (2009) "Firm-Specific Human Capital: A Skill-Weights Approach," Journal of Political Economy, 117 (5), 914–940.
- Lee, David S., Justin McCrary, Marcelo J. Moreira, and Jack R. Porter (2021) "Valid t-ratio Inference for IV," Working Paper 29124, National Bureau of Economic Research.
- Lise, Jeremy and Jean-Marc Robin (2017) "The Macrodynamics of Sorting between Workers and Firms," *The American Economic Review*, 107 (4), 1104–1135.
- Maestas, Nicole, Kathleen J. Mullen, David Powell, Till von Wachter, and Jeffrey B. Wenger (2018) "The Value of Working Conditions in the United States and Implications for the Structure of Wages," Working Paper 25204, National Bureau of Economic Research.
- Manski, Charles F. (1993) "Identification of Endogenous Social Effects: The Reflection Problem," The Review of Economic Studies, 60 (3), 531–542.

- Mas, Alexandre and Amanda Pallais (2019) "Labor Supply and the Value of Non-Work Time: Experimental Estimates from the Field," *American Economic Review: Insights*, 1 (1), 111–126.
- McCue, Kristin (1996) "Promotions and Wage Growth," Journal of Labor Economics, 14 (2), 175–209.
- Menezes-Filho, Naércio Aquino, Marc-Andreas Muendler, and Garey Ramey (2008) "The Structure of Worker Compensation in Brazil, with a Comparison to France and the United States," *Review* of Economics and Statistics, 90 (2), 324–346.
- Milgrom, Paul and John Roberts (1992) *Economics, Organization and Management*, Englewood Cliffs, N.J: Pearson, 1st edition.
- Mincer, Jacob A. (1974) Schooling, Experience, and Earnings: NBER.
- Moscarini, Giuseppe and Fabien Postel-Vinay (2018) "The Cyclical Job Ladder," Annual Review of Economics, 10 (1), 165–188.
- Murphy, Kevin J. (1985) "Corporate performance and managerial remuneration: An empirical analysis," *Journal of Accounting and Economics*, 7 (1), 11–42.
- Newey, Whitney K. and Daniel McFadden (1994) "Chapter 36 Large sample estimation and hypothesis testing," in *Handbook of Econometrics*, 4, 2111–2245: Elsevier.
- OECD (2019) "Higher education in Brazil," in Rethinking Quality Assurance for Higher Education in Brazil, 65–90: OECD.
- OECD-IDB (2014) "Detailed description of Employment Protection Legislation in LAC countries."
- O'Flaherty, Brendan and Aloysius Siow (1995) "Up-or-Out Rules in the Market for Lawyers," Journal of Labor Economics, 13 (4), 709–735.
- Oreopoulos, Philip, Till von Wachter, and Andrew Heisz (2012) "The Short- and Long-Term Career Effects of Graduating in a Recession," *American Economic Journal: Applied Economics*, 4 (1), 1–29.
- Pastorino, Elena (2019) "Careers in Firms: the Role of Learning and Human Capital."
- Perry, Guillermo E., Omar Arias, Pablo Fajnzylber, William F. Maloney, Andrew Mason, and Jaime Saavedra-Chanduvi (2007) Informality: Exit and Exclusion: The World Bank.
- Postel–Vinay, Fabien and Jean-Marc Robin (2002) "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70 (6), 2295–2350.
- Prendergast, Canice (1993) "The Role of Promotion in Inducing Specific Human Capital Acquisition," The Quarterly Journal of Economics, 108 (2), 523–534.
- Schmieder, Johannes F., Till von Wachter, and Joerg Heining (2020) "The Costs of Job Displacement over the Business Cycle and Its Sources: Evidence from Germany."
- Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter (2019) "Firming Up Inequality," *The Quarterly Journal of Economics*, 134 (1), 1–50.
- Sorkin, Isaac (2018) "Ranking Firms Using Revealed Preference," The Quarterly Journal of Economics.
- Sun, Liyang and Sarah Abraham (2021) "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 225 (2), 175–199.
- Taber, Christopher and Rune Vejlin (2020) "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market," *Econometrica*, 88 (3), 1031–1069.

- Topel, Robert H. and Michael P. Ward (1992) "Job Mobility and the Careers of Young Men," *The Quarterly Journal of Economics*, 107 (2), 439–479.
- Waldman, Michael (1984) "Job Assignments, Signalling, and Efficiency," The RAND Journal of Economics, 15 (2), 255–267.
Tables and Figures



Figure 1: Example of job output functions

Note: The figure shows an example of two job output functions that generate a motive for promotions. Workers who are believed to be of high ability have higher expected output in the high risk occupation (job 2), whereas workers who are believed to be of low ability have higher expected outcomes in the low risk occupation (job 1). Both occupations provide negative output when workers are sufficiently likely to be of low ability, so the outside option is preferred.







Figure 3: Prevalence of Promotions for Production Workers

Note: The figure plots the share of workers each year that advanced from an elementary production occupation in the previous year to a supervisor occupation in the current year. The "within occupation" series restricts the definition of promotions to purely vertical moves that are within the same occupational group. The sample is all workers in the Brazilian formal sector between the ages of 25 and 50. For more details, see Section 3.3.

Figure 4: Timeline of empirical approach



Note: The figure describes the years of data used in each section of the empirical approach. The first four years of data (2003-2006) are reserved for estimating firms' promotion propensities (see Section 4.1). The analysis sample consists of workers who moved firms between 2008 and 2012 and have three years of pre-move data (see Section 4.2). My data end in 2015, so the length of post-move outcomes ranges from 4 years (for the 2012 cohort) to 8 years (for the 2008 cohort).



Figure 5: Test for movers balance

Note: The figure plots the difference in average worker characteristics between those that moved to high promotion firms and those that moved to low promotion firms. Age and years of schooling are both scaled by their overall means, so the coefficients are interpreted as relative differences. The first set of estimates is the raw difference across all worker cohorts and only controls for possible differences in composition across worker cohorts. The second set of estimates adds state-by-occupation-by-cohort fixed effects. The final set of estimates also controls for the estimated firm wage premium at the worker's origin firm. Standard errors are clustered by the worker's destination firm, and lines indicate the 95% confidence interval.



Figure 6: Impact of high promotion firms on worker promotions

Note: The figure plots event study coefficients from estimating Equation 6 on the baseline sample of all promotion track workers between the ages of 25 and 50 that made an employer-to-employer transition to a high or low promotion firm. The coefficients are the estimated yearly differences in the outcomes of workers who moved to a high promotion firm relative to workers who moved to a low promotion firm, averaged across all five cohorts of movers. The outcome is whether the worker is working as a supervisor and includes workers who have left formal employment (assumed to not be working as supervisors). Standard errors are clustered by the worker's destination firm, and lines indicate the 95% confidence interval.



Figure 7: Impact of high promotion firms on formal labor market attachment

Note: The figure plots event study coefficients from estimating Equation 6 on the baseline sample of all promotion track workers between the ages of 25 and 50 that made an employer-to-employer transition to a high or low promotion firm. The coefficients are the estimated yearly differences in the outcomes of workers who moved to a high promotion firm relative to workers who moved to a low promotion firm, averaged across all five cohorts of movers. The outcome is whether a worker is working in formal employment. Standard errors are clustered by the worker's destination firm, and lines indicate the 95% confidence interval.



Figure 8: Impact of high promotion firms on log earnings for employed workers

Note: The figure plots event study coefficients from estimating Equation 6 on the baseline sample of all promotion track workers between the ages of 25 and 50 that made an employer-to-employer transition to a high or low promotion firm. The coefficients are the estimated yearly differences in the outcomes of workers who moved to a high promotion firm relative to workers who moved to a low promotion firm, averaged across all five cohorts of movers. The outcome is log earnings and is only defined for workers in formal employment. Standard errors are clustered by the worker's destination firm, and lines indicate the 95% confidence interval.

Figure 9: Robustness of main estimates to controls





Note: The figures compare pooled event study coefficients from estimating variations of Equation 6 on the baseline sample. The top panel plots the pooled effects within two years of moving, and the bottom panel plots the pooled effects more than two years after the move. The outcomes are whether the worker is working in a supervisor position, whether the worker is working in any job in formal employment, and the worker's log earnings conditional on being in formal employment. Basic controls only include age, gender, education controls, and two-way worker and year fixed effects. Extended controls include all of the baseline controls, and additionally include flexible trends based on workers' baseline earnings as well as time-varying changes to the gender and educational wage structure. Standard errors are clustered by the worker's destination firm, and lines indicate the 95% confidence interval.



Figure 10: Comparison of main effects between identification strategies

Note: The figures compare pooled event study estimates from the baseline research design in Section 4.2 to estimates from the extensions in Section 4.3 that incorporate additional sources of variation. The top panel plots the pooled effects within two years of moving, and the bottom panel plots the pooled effects more than two years after the move. The outcomes are whether the worker is working in a supervisor position, whether the worker is working in any job in formal employment, and the worker's log earnings conditional on being in formal employment. The mass layoff estimates restrict the sample to workers whose origin firms experienced mass layoff events at the time of the workers' moves. The local hiring estimates are the instrumental variables estimates from using the jack-knifed municipal hiring share as an instrument for the worker's destination firm. Standard errors are clustered by the worker's destination firm for the mass layoff estimates and by the destination municipality for the local hiring estimates, and lines indicate the 95% confidence interval.



Figure 11: Differential impact of high promotion firms by worker job levels

Note: The figure plots event study coefficients from estimating Equation 9 on the baseline sample. Workers are split into four groups based on the type of their destination firm as well as whether they were promoted within a year of moving. The plotted groups are high promotion firm movers who were not promoted, high promotion firm movers who were promoted, and low promotion firm movers who were promoted, respectively. The reference group is low promotion firm movers who were not promoted. The outcome is log earnings conditional on being in formal employment. Standard errors are clustered by the worker's destination firm, and lines indicate the 95% confidence interval.





Note: The figure decomposes the baseline estimate of the effects of high promotion firms on formal labor market attachment by estimating Equation 6 on each type of worker separation. The coefficients are the estimated difference in the outcomes of workers who moved to a high promotion firm relative to workers who moved to a low promotion firm, averaged across all five cohorts of movers. The first outcome is a spell outside formal employment following any recorded reason for separating from their previous employer. The second outcome is a spell outside formal employment following an employer-initiated separation. The third outcome is a spell outside formal employment following a worker-initiated separation. Standard errors are clustered by the worker's destination firm, and lines indicate the 95% confidence interval.



Figure 13: Comparison of main effects by worker potential

Note: The figures compare pooled event study coefficients from separately estimating Equation 6 on high and low potential workers from the baseline sample. The top panel plots the pooled effects within two years of moving, and the bottom panel plots the pooled effects more than two years after the move. The outcomes are whether the worker is working in a supervisor position, whether the worker is working in any job in formal employment, and the worker's log earnings conditional on being in formal employment. Worker potential is defined as the worker's predicted likelihood of being promoted within a year after moving firms based on their characteristics before moving firms (estimated from Equation 10 using a holdout sample of workers). High potential workers are workers in the top tercile of worker potential, and low potential workers are workers in the bottom tercile. Standard errors are clustered by the worker's destination firm, and lines indicate the 95% confidence interval.



Figure 14: Correlation in firm practices

Note: The figure plots the correlation between the plant's willingness to promote high performing workers and the plant's willingness to fire low performing workers, as scored from structured interviews from the World Management Survey. Scores range from 1 to 5 on each question, with 5 as reflecting the most active worker management practices and 1 as reflecting the least active. For more details, see Section 6.4.

Occupation Code	Title
7601XX	Supervisors of the textile industry
7610XX	Multipurpose workers in the textile industry
7611XX	Workers in the classification and washing of fibers
7612XX	Operators of spinning machines
7613XX	Operators of looms
7614XX	Workers in finishing, dying, and stamping
7618XX	Inspectors and reviewers of textile production

Table 1: Example of occupational group with observable line of progression

Table 2: Estimates of promotion wage premia						
	Outcome: Log Earnings					
	(1)	(2)	(3)	(4)	(5)	
Supervisor	0.628***	0.520***	0.384***	0.170***	0.200***	
	(0.0180)	(0.0185)	(0.0191)	(0.00347)	(0.00309)	
Number of observations	11581373	11581278	11581208	11092695	5616458	
Number of workers	2473852	2473850	2473846	1985349	1071791	
Adjusted R^2	0.234	0.407	0.520	0.890	0.874	
Controls:						
State-occupation-year FE		Υ	Υ	Υ	Υ	
Worker controls			Υ	Υ	Υ	
Worker FE				Y	Υ	
Firm wage premia					Υ	

T-11. 9. E-4: c

Note: The table reports estimates for the average supervisor wage premium in the Brazilian formal employment sector. The sample is a 5% sample of all workers between the ages of 25 and 50 that work in occupational groups with a well-defined supervisor role. The first column reports the bivariate wage premium and controls only for year fixed effects. Worker controls are a quadratic in age interacted with gender along with indicators for the worker's education, race, and state. Firm wage premia controls net out the estimated AKM firm fixed effect from worker earnings. Standard errors are two-way clustered at the worker and firm level. Stars indicate the level of significance: * 5%, ** 1%, and *** 0.1%.

	Outcome: Log Earnings					
	(1)	(2)	(3)	(4)	(5)	(6)
Age	0.0173***	0.0163***	0.0115***	0.0105***	0.0133***	0.0118***
	(0.000294)	(0.000321)	(0.000185)	(0.000195)	(0.000238)	(0.000258)
Supervisor		0.412***		0.393***		0.434***
		(0.0130)		(0.0112)		(0.0135)
Share of age		0.0547		0.0880		0.111
coefficient						
explained						
Number of obs.	6056674	6056674	6056525	6056525	3217367	3217367
Number of workers	1697830	1697830	1697814	1697814	978304	978304
Adjusted \mathbb{R}^2	0.474	0.489	0.538	0.551	0.483	0.504
Controls:						
Worker controls	Υ	Υ	Y	Υ	Y	Y
Occupation			Υ	Y	Υ	Υ
controls						
Firm wage premia					Υ	Y

Note: The table reports estimates of the average lifecycle wage profile with and without controlling for the average supervisor wage premium in the Brazilian formal employment sector. The sample is a 5% sample of all workers between the ages of 25 and 35 that work in occupational groups with a well-defined supervisor role. The share of age coefficient explained is the relative decrease in the age coefficient after adding an indicator for whether the worker is working as supervisor. Worker controls are indicators for gender, worker education, and race as well as state-by-year fixed effects. Occupational controls are the worker's occupational tenure and state-by-occupational-group-by-year fixed effects. Firm wage premia controls net out the estimated AKM firm fixed effect from worker earnings. Standard errors are two-way clustered at the worker and firm level. Stars indicate the level of significance: * 5%, ** 1%, and *** 0.1%.

		Standard	Low prom.	
Outcome:	Difference	errors	firm mean	Adj. R^2
Promotions	0.0172^{***}	(0.000152)	0.000159	0.354
Exit Formal Employment	0.0165^{***}	(0.000744)	0.165	0.00740
Log Employment	0.490***	(0.00934)	3.584	0.0415
Age	-0.310***	(0.0186)	35.27	0.00376
Log Earnings	0.0900***	(0.00417)	6.462	0.00677
AKM Firm FE	0.0542^{***}	(0.00346)	-	0.00443
Supervisor Share	0.0392^{***}	(0.000612)	0.0302	0.0789
Log Earnings Growth	0.00990^{***}	(0.000424)	-	0.00869
Log Employment Growth	0.0338***	(0.00162)	0.0355	0.00746
Number of firms		694	417	

Table 4: Differences between high and low promotion firms

Note: The table reports the difference between high and low promotion firms in average firm characteristics between 2004-2006, as well as the means for low promotion firms in each relevant category. The adjusted R^2 reports the share of firm-level variation in each outcome that is explained by the firm's high versus low promotion status. Stars indicate the level of significance: * 5%, ** 1%, and *** 0.1%.

	Baseline	Mass layoffs
Number of workers	1100590	127058
Number of origin firms	162573	16103
Number of destination firms	49785	21170
Worker characteristics before move:		
Age	32.94	33.89
	[6.761]	[7.007]
Female	0.262	0.241
	[0.439]	[0.428]
Years of schooling	10.50	9.866
	[2.941]	[3.033]
Monthly earnings (2010 Reals)	1242.9	1186.4
	[900.4]	[831.7]
Share to high promotion firm	0.417	0.442
Worker outcomes > 2 years following move:		
In formal employment	0.777	0.764
At destination firm	0.305	0.278
In supervisor occupation	0.0280	0.0284

Table 5: Analysis sample summary statistics

Note: The table reports summary statistics about the job-to-job movers in the baseline sample and the mass-layoffs subsample. Pre-move worker characteristics refer to the snapshot of worker data from the year before the job-to-job transition (t = b - 1). Outcomes more than two years following the move are averaged over all relevant years where the data are available. All statistics on characteristics and outcomes are in means, and standard deviations for continuous measures are in brackets.

	Outcome: Log Earnings				
	(1)	(2)	(3)	(4)	(5)
High Firm Share (β_1)	0.132***	0.140***	0.140***	0.0397^{**}	0.00850
	(0.0162)	(0.0151)	(0.0152)	(0.0134)	(0.00544)
High Firm Share \times Super. (β_2)	0.244^{***}	0.237^{***}	0.196^{***}	0.106^{***}	0.159^{***}
	(0.0423)	(0.0405)	(0.0334)	(0.0287)	(0.0306)
Number of observations	11569458	11569388	11569388	11569388	11187050
Number of municipalities	5500	5500	5500	5500	5499
Adjusted R^2	0.408	0.521	0.523	0.539	0.771
Controls:					
State-occupation-year FE	Υ	Υ	Υ	Υ	Y
Worker controls		Y	Y	Y	Y
Occupation-superyear FEs			Υ	Υ	Y
Municipality controls				Υ	
Firm FEs					Υ

Table 6: Estimates of regional wage premia

Note: The table reports estimates for municipal differences in wages and supervisor wage premia in the Brazilian formal employment sector (estimated using Equation 11). The sample is a 5% sample of all workers between the ages of 25 and 50 that work in occupational groups with a well-defined supervisor track in municipalities with at least 100 workers. The estimands of interest are the wage premium for all workers in municipalities with a high share of high promotion employers (β_1) as well as the differential wage premium for promoted workers in municipalities with a high share of high promotion employers (β_2). Worker controls are a quadratic in age interacted with gender along with indicators for the worker's education, race, and state. Municipality controls are controls for log employment and the overall hiring share of either high or low promotion firms that enter both linearly and as interactions with supervisor status. Standard errors are clustered by municipality. Stars indicate the level of significance: * 5%, ** 1%, and *** 0.1%.

	Description	Moments	Assumptions
Calib	rated parameters:		
κ	Degree of asymmetric information	Share of job changers that remain supervisors	Calibrated from movers data
δ	Prob. of exogenous separation	Estimated job destruction rate	From Dix-Carneiro et al. (2021)
Jointl	y estimated parameters:		
\bar{Q}_H	Rate of learning for high promotion firms	Average share promoted	Constrained efficient promotions $+$ firings
\bar{Q}_L	Rate of learning for low promotion firms	Average share in formal labor market attachment	Random labor market matching
α	Share of high ability workers	Effect of high prom. firms on promotions	
g	Prob. of re-employment	Effect of high prom. firms on	
	upon separation	formal labor market attachment	

Table 7: Summary of parameters and moments for identification

-.

Note: The table summarizes the moments and assumptions required to identify the parameters in the partial equilibrium model of employer learning and job assignment. Two parameters - the degree of asymmetric information κ and the probability of exogenous separation δ are directly calibrated. The remaining parameters are jointly identified from the means and treatment effects from the baseline movers analysis. The identifying assumptions and model restrictions are discussed in more detail in Section 8.1.

Definition	Description	Estimate	S.E.
Parameters:			
κ	Degree of asymmetric information	0.300	-
δ	Prob. of exogenous separation	0.282	-
$ar{Q}_H$	Rate of learning for high promotion firms	0.208	(0.029)
$ar{Q}_L$	Rate of learning for low promotion firms	0.150	(0.022)
α	Share of high ability workers	0.214	(0.031)
g	Prob. of re-employment upon separation	0.414	(0.028)
Implied measur	es:		
$\left(\alpha - \alpha' \right) / \alpha$	Degree of adverse selection in job movers	0.292	(0.039)
$\left(\alpha - E\left[P\right]\right)/\alpha$	Share of high ability workers not in supervisor role	0.869	(0.019)

Table 8: Baseline model estimates

Note: The table reports the estimates from quantifying the partial equilibrium model of employer learning and job assignment. The quantification approach is described in Section 8.1. The definition of high and low promotion firms match the classification used for the estimation of treatment effects, as described in Section 4.1, and the resulting estimates quantify the difference in economic behavior between the two groups. The degree of adverse selection in job movers is the relative decrease in the likelihood that a job mover is of high ability as compared to the general population. Meanwhile, the share of high ability workers that are not working as supervisors reflects skill misallocation. Standard errors are calculated from the joint variance-covariance matrix of empirical moments, which are clustered by the worker's destination firm.

A Additional Tables and Figures



Figure A.1: Prevalence of Occupations with Clear Promotion Tracks



Figure A.2: Distribution of firm promotion propensities

Note: The figure plots a kernel density estimate of the distribution of firm promotion propensities (η_j) between 2004 and 2006. Each firm's promotion propensity is calculated as the residual firm promotion rates after controlling for differences in worker characteristics and occupational groups and averaged over the three years. Estimates in the figure have been winsorized at the 5th and 95th percentiles. For more details, see Section 4.1.



Figure A.3: Age profile of supervisors and promotions



Figure A.4: Visual intuition for IV first stage

Figure A.5: Impact of high promotion firms on a worker ever being promoted





Figure A.6: Impact of high promotion firms on worker promotions (by cohort)

Figure A.7: Impact of high promotion firms on formal labor market attachment (by cohort)





Figure A.8: Impact of high promotion firms on log earnings for employed workers (by cohort)

Figure A.9: Impact of high promotion firms on worker promotions (continuous measure)



Effect of High Opp. Firm: Promotion (w/ Exits)



Figure A.10: Impact of high promotion firms on formal labor market attachment (continuous measure)

Figure A.11: Impact of high promotion firms on log earnings for employed workers (continuous measure)





Figure A.12: Impact of high promotion firms on worker promotions (mass layoffs)

Figure A.13: Impact of high promotion firms on formal labor market attachment (mass layoffs)





Figure A.14: Impact of high promotion firms on log earnings for employed workers (mass layoffs)

Figure A.15: Impact of high promotion firms on worker promotions (local hiring IV)



63



Figure A.16: Impact of high promotion firms on formal labor market attachment (local hiring IV)

Figure A.17: Impact of high promotion firms on log earnings for employed workers (local hiring IV)





Figure A.18: Comparison of main effects between alternative firm types

• Baseline 🛛 Internal Promotions 🔹 Wage Growth



Figure A.19: Robustness of model estimates to alternate assumptions

(a) Robustness to alternate kappa



(b) Robustness to alternate delta

Note: The figure shows the model estimates under alternative assumed values on κ , the degree of asymmetric information on the secondary market, and on δ , the exogenous job separation rate. Adverse selection refers to the relative share of high ability workers in the pool of job-movers compared to the population share of high ability workers. Misallocation refers to the likelihood that a high ability worker is not working as a supervisor. The model can no longer exactly fit the sample means and treatment effect estimates at 50% of the baseline assumed value for δ , so those estimates should be interpreted with some caution. The estimates under the baseline assumed value are indicated by dashed lines.

Measure	Correlation Coefficient				
Baseline measure	1				
Stayers only	0.925	1			
No controls	0.947	0.885	1		
All promotions	0.731	0.673	0.700	1	

Table A.1: Correlation between measures of firm promotions

Note: This table reports the correlation matrix between the baseline measure of the firm's promotion propensity η_j and alternate measures. All measures reflect the firm's average η_{jt} over 2004-2006 and are winsorized at the 5th and 95th percentiles. The "stayers only" measure restricts the sample to workers who remained in the same firm, so the estimate reflects firms' internal promotion rates. The "no controls" measure removes any controls and only considers the firms' raw promotion rate. "All promotions" expands the baseline sample to also include workers who changed occupation groups.

Table A.2: Pooled estimates on the impact of high promotion firms on other worker outcomes

	(1)	(2)	(3)	(4)	(5)
	Employed	Promoted			
	at	(cond. on	Earnings		Earnings
	destination	formal	net of firm	Contractual	(adjusted
Outcome:	firm	emp.)	wage premia	salary	for U.E.)
Effects within 2 years of	-0.0263***	0.0136^{***}	-0.00625	0.00964	0.00799
moving	(0.0041)	(0.00097)	(0.0036)	(0.0071)	(0.0055)
Effects more than 2	-0.0447***	0.0132^{***}	-0.00745	0.0119^{*}	-0.00665
years after moving	(0.0082)	(0.0010)	(0.0043)	(0.0060)	(0.0058)
Number of observations	8279318	7378436	6640880	6390190	8279252
Number of dest. firms	45870	45870	45654	45496	45870
Adjusted \mathbb{R}^2	0.617	0.304	0.795	0.776	0.775

Note: The table reports additional pooled event study coefficients from estimating Equation 6 on the baseline sample. The outcomes are whether the worker is employed at their original destination firm, whether the worker is promoted (conditional on the worker remaining in formal employment), the worker's log earnings after netting out the AKM firm effect, the worker's contractual salary, and a lower bound for worker earnings that include exiters (adjusting for the value of non-work time). Standard errors are clustered by the worker's destination firm. Stars indicate the level of significance: *5%, **1%, and ***0.1%.

	Baseline	High	Low
		Potential	Potential
Number of workers	1100590	198732	169589
Number of origin firms	162573	25758	26544
Number of destination firms	49785	26223	26347
Worker characteristics before move:			
Age	32.94	33.11	32.56
	[6.761]	[6.613]	[6.774]
Female	0.262	0.284	0.234
	[0.439]	[0.451]	[0.423]
Years of schooling	10.50	11.81	9.191
	[2.941]	[2.716]	[2.966]
Monthly earnings (2010 Reals)	1242.9	1869.9	874.5
	[900.4]	[1251.4]	[328.1]
Share to high promotion firm	0.417	0.435	0.406
Worker outcomes > 2 years following move:			
In formal employment	0.777	0.795	0.759
At destination firm	0.305	0.326	0.286
In supervisor occupation	0.0280	0.0476	0.0125

Table A.3: Worker potential subsample summary statistics

Note: The table reports summary statistics about the job-to-job movers in the baseline sample and the high and low potential worker subsamples. Pre-move worker characteristics refer to the snapshot of worker data from the year before the job-to-job transition (t = b - 1). Outcomes more than two years following the move are averaged over all relevant years where the data are available. All statistics on characteristics and outcomes are in means, and standard deviations for continuous measures are in brackets.

Question	Score	Criteria
Developing Talent and	1	People are promoted primarily upon the basis of
Promoting High		tenure
Performers	5	We actively identify, develop and promote our
		top performers
Removing Poor	1	Poor performers are rarely removed from their
Performers/Making		positions
Room for Talent	5	We move poor performers out of the company or
		to less critical roles as soon as a weakness is
		identified

Table A.4: World Management Survey questions and scoring criteria

Note: The table reproduces the question and scoring rubric for the promotion and firing practice questions from the World Management Survey. All survey responses range from 1 to 5, with 5 indicating the most active firm practice and 1 indicating the least active. For more details, see Section 6.4.

	Outcome: Leave formal labor market (next year)						
	(1)	(2)	(3)	(4)	(5)		
High Firm Share (β_1)	0.0294^{***}	0.0154^{***}	0.0155^{***}	0.0213^{***}	0.00779^{*}		
	(0.00315)	(0.00298)	(0.00298)	(0.00296)	(0.00304)		
High Finns Chang & Curron (2)	0.0205***	0 0200***	0.0005***	0.0176**	0.0957***		
High Firm Share × Super. (p_2)	-0.0383	-0.0382	-0.0285	-0.0170	-0.0557		
	(0.00620)	(0.00613)	(0.00574)	(0.00567)	(0.00602)		
Number of observations	10489980	10489910	10489910	10489910	10127824		
Number of municipalities	5489	5489	5489	5489	5487		
Adjusted R^2	0.0233	0.0368	0.0371	0.0373	0.113		
Controls:							
State-occupation-year FE	Y	Y	Y	Υ	Y		
Worker controls		Y	Υ	Υ	Υ		
Occupation-superyear FEs			Y	Υ	Y		
Municipality controls				Υ			
Firm FEs					Y		

Table A.5: Estimates of regional differences in labor market attachment

Note: The table reports estimates for municipal differences formal labor market attachment in the Brazilian formal employment sector (estimated using Equation 11). The sample is a 5% sample of all workers between the ages of 25 and 50 that work in occupational groups with a well-defined supervisor track in municipalities with at least 100 workers. The estimands of interest are the likelihood of leaving formal employment for all workers in municipalities with a high share of high promotion employers (β_1) as well as the differential likelihood for promoted workers in municipalities with a high share of high promotion employers (β_2). Worker controls are a quadratic in age interacted with gender along with indicators for the worker's education, race, and state. Municipality controls are controls for log employment and the overall hiring share of either high or low promotion firms that enter both linearly and as interactions with supervisor status. Standard errors are clustered by municipality. Stars indicate the level of significance: * 5%, ** 1%, and *** 0.1%.

Parameter	Sensitivity to η							
$ar{Q}_H$	8.13	0.95	-19.32	-11.14				
$ar{Q}_L$	7.58	0.68	-19.02	-7.99				
α	-2.23	-1.08	24.59	11.96				
g	7.31	3.55	-21.58	-10.49				

Table A.6: Sensitivity of model parameters to empirical moments (Λ)

Note: The ordering of the empirical moments are the average share of workers promoted, the average share of workers in formal employment, the effect of high promotion firms on worker promotions, and the effect of high promotion firms on formal employment. The interpretation of the sensitivity matrix is that for a local perturbation of the empirical moments that converges to η , the first-order asymptotic bias in the parameter estimates are $\Lambda \eta$. For more details, see Section C.3.2.

B Proofs of Propositions

B.1 Partial Equilibrium Model

Proposition. If information about job assignments on the secondary market is sufficiently weak (so κ is sufficiently small), then a unique Perfect Bayesian equilibrium exists. In this equilibrium, (i) job assignments for workers are efficient (given firms' information about workers) (ii) all turnover is involuntary (iii) wages are given by Equations 2 and 4.

Proof. I characterize the equilibrium by solving the problem backward. I conjecture that the incumbent firms will promote workers that are revealed to be high ability, fire workers that are revealed to be low ability, and retain all workers whose abilities are still unobserved, and I show that this conjectured behavior is optimal given secondary market wages. I then prove uniqueness by showing that this conjectured behavior is the only one consistent with an equilibrium.

For separated workers who successfully convinced the secondary market that they were previously promoted, the secondary market will infer that they were high ability and offer the expected output

$$w_2^S = f_2\left(\theta_H\right).$$

Meanwhile, the likelihood that a low ability worker enters the unknown secondary market workers pool is

$$\delta + (1 - \delta) \left(\underbrace{\rho Q_H + (1 - \rho) Q_L}_{\text{prob. fired}} \right) \equiv \delta + (1 - \delta) \bar{Q},$$

whereas the likelihood that a high ability worker enters the unknown worker pool is $\delta (1 - \bar{Q}\kappa)$. By Bayes' rule,

$$\begin{aligned} \alpha' &\equiv Pr(\theta = \theta_H | \text{unknown quality}) \\ &= \frac{\alpha \delta \left(1 - \bar{Q}\kappa\right)}{\alpha \delta \left(1 - \bar{Q}\kappa\right) + (1 - \alpha) \left(\delta + (1 - \delta) \bar{Q}\right)} \\ &< \alpha. \end{aligned}$$

Since low ability workers are both more likely to be fired and less likely to convince new employers that they are high ability, they are going to comprise a higher share of the secondary market workers of unknown quality than in the general population. So the secondary market is adversely selected. The degree of adverse selection is higher when there is greater learning by initial employers (\bar{Q}) or by the secondary market (κ), and lower when there are more exogenous separations (δ). Assuming that it is still optimal for secondary market firms to assign workers of unknown quality to the low complexity occupation, offered wages for these workers will be their expected output, so

$$w_1^S = \alpha' f_1(\theta_H) + (1 - \alpha') f_1(\theta_L).$$

Next, I can then verify that the conjectured job assignment by incumbent firms is indeed optimal given the secondary market offers w_1^S, w_2^S and characterize the incumbent firms' wage offers.

The incumbent firms earn positive profits from promoted workers of known high ability regardless
of g because the worker is not guaranteed to remain promoted in the secondary market. So, the incumbent firm needs to offer a promoted worker only their expected outside wage to retain them:

$$w_2^I = g\left[\kappa w_2^S + (1-\kappa)w_1^S\right] < gf_2\left(\theta_H\right).$$

Similarly, the incumbent firm also benefits by retaining the workers whose ability it did *not* observe since the firm has private information that those workers are not adversely selected. For those workers, the probability that they are the high type is still α , so the firm makes positive profits by retaining those unknown workers and offering them the expected outside wage for workers of unknown ability:

$$w_1^I = g \left[\alpha' f_1 \left(\theta_H \right) + \left(1 - \alpha' \right) f_1 \left(\theta_L \right) \right] < g E \left[f_1 \left(\theta \right) |\alpha| \right]$$

Finally, for workers who are observed to be low ability, $f_2(\theta_L) < f_1(\theta_L) < 0$, so the firm is better off firing them rather than retaining them with any positive wage.

For this equilibrium to exist, two conditions need to hold. First, the expected productivity for unknown workers in the secondary market needs to be positive, so

$$\underbrace{\alpha' f_1(\theta_H) + (1 - \alpha') f_1(\theta_L)}_{=E[f_1(\theta)|\alpha']} \ge 0.$$

This condition also implies that the incumbent firm will also find it optimal to retain workers whose ability is unobserved, since $\alpha > \alpha'$. In addition, the wage gain upon promotion needs to be sufficiently low for the incumbent firm to still prefer to promote workers of high ability, so

$$f_2(\theta_H) - w_2^I \ge f_1(\theta_H) - w_1^I.$$

Rearranging using the expression for wages,

$$f_{2}(\theta_{H}) - f_{1}(\theta_{H}) \geq w_{2}^{I} - w_{1}^{I}$$
$$= g \left[\kappa w_{2}^{S} + (1 - \kappa) w_{1}^{S} - w_{1}^{S} \right]$$
$$= g \kappa \left(f_{2}(\theta_{H}) - E \left[f_{1}(\theta) | \alpha' \right] \right)$$

 α' is decreasing in κ (since increasing the informativeness of job assignments reduces the likelihood that a high ability worker joins the unknown workers pool), so the right hand side of the inequality is also decreasing in κ . So, decreasing κ relaxes both inequality constraints, and both conditions clearly hold when $\kappa \to 0$ as long as

$$E\left[f_{1}\left(\theta\right)|\alpha'=\alpha\delta/\left(\delta+\left(1-\alpha\right)\left(1-\delta\right)\bar{Q}\right)\right]\geq0,$$

so the equilibrium exists whenever κ is sufficiently small.

Finally, I show that if κ is sufficiently small such that the equilibrium exists, it is also the unique equilibrium. To do so, it suffices to show that for the incumbent firm, firing low ability workers, promoting high ability workers, and retaining unknown workers is the only strategy that is consistent with an equilibrium.

Clearly, there can be no equilibrium that exists where the incumbent firm retains all workers

since the output of workers with $\theta = \theta_L$ is strictly less than the outside option. The firm will also never retain low ability workers but fire unknown workers, since the market wages for those two types are identical and the expected productivity of the latter strictly dominates the former.

Similarly, there can be no equilibrium where the incumbent firm retains only high ability workers. Suppose the incumbent firms retain high ability workers and fire all other workers. This implies that expected outside wages exceed expected productivity for workers whose ability is unknown, which cannot be consistent with the competitive secondary market since the secondary market is adversely selected.

So, it suffices to consider whether an equilibrium exists where the incumbent firm treats high ability workers differently. There can be no equilibrium where the incumbent firm will fire high ability workers, since their expected outside wage will always be weakly less than $f_2(\theta_H)$. Similarly, if the incumbent firm keeps all high ability workers in the low complexity job, secondary market wages for workers of unknown quality, w_1^S , rises, while secondary market wages for any promoted workers would remain the same. This strictly relaxes the incentive compatibility problem for promotions, so the incumbent firm would profitably deviate by promoting high ability workers instead.

It then follows that under the conditions where the conjectured equilibrium exists, the equilibrium where the incumbent firm promotes high ability workers, fires low ability workers, and retains workers of unknown ability in the low complexity job is also the unique equilibrium.

Proposition. In the equilibrium described in Proposition 1, workers initially employed at high learning firms are (i) more likely to be promoted and (ii) more likely to become unemployed than workers initially employed at low learning firms.

Proof. This follows directly from the characterization of the equilibrium. The incumbent firm will promote all high ability workers and fire all low ability workers. Since the composition of workers and labor market parameters are the same between high and low learning firms, the results follow from differences in learning rates.

Specifically, the likelihood that a worker starting at firm f will be promoted by the end of the period is

$$\Pr\left(\operatorname{Promoted}|Q_f\right) = \alpha Q_f \left(1 - \delta + \delta g \kappa\right),$$

and the likelihood that a worker will be be unemployed is

$$\Pr\left(\text{Unemployed}|Q_f\right) = (1-g)\left[\delta + (1-\alpha)\left(1-\delta\right)Q_f\right].$$

Both likelihoods are increasing in Q_f .

B.2 General Equilibrium Model

Proposition. Suppose (i) the marginal cost for vacancies c'(v) is log-concave and (ii) the elasticity of re-employment with respect to the number of high learning employer vacancies exceed the elasticity of expected secondary market worker output. Then, in more productive regions: (i) high learning employers post a greater share of vacancies, (ii) offered wages for incumbent workers are higher (iii) wage differentials between promoted and unpromoted workers are larger.

Proof. Under the assumptions on labor market matching, the market level of high learning firm vacancies V_H and low learning firm vacancies V_L imply the labor market parameters ρ and g as:

$$\rho = \frac{V_H}{V_H + V_L}$$
$$g = \frac{M\left(m^I, V_H + V_L - m^I\right)}{m^I},$$

where M(l, v) is the reduced form matching function that relates the level of job matches to the number of workers l and vacancies v, and $m^{I} = (l, V_{H} + V_{L})$ is the number of initially matched workers. Conditional on the number of vacancies (and their implied labor market parameters), the job assignment, offer, and secondary market decisions will follow the characterization in Proposition 1. So, the firm's expected profits from a vacancy filled with a new worker is

$$E\left[\pi_{f}|V_{H},V_{L}\right] = \left(1-\delta\right) \left\{ Q_{f}\alpha \left(\underbrace{f_{2}\left(\theta_{H}\right)-w_{2}^{I}}_{\text{profits from promoted}}\right) + \left(1-Q_{f}\right) \left(\underbrace{E\left[f_{1}\left(\theta\right)|\alpha\right]-w_{1}^{I}}_{\text{profits from unpromoted}}\right) \right\},$$

where V_H , V_L are the total high and low learning firm vacancies in the market, respectively, and all other objects match their definition from Proposition 1. The firm profits only upon successfully retaining an initially assigned worker. Any possibility that a vacancy is filled in the secondary market is irrelevant since the secondary market is competitive.

In the vacancy creation problem, the firm solves

$$\max_{v} h\left(V_{H}, V_{L}\right) E\left[\pi_{f} | V_{H}, V_{L}\right] v - c\left(v\right),$$

where $h(V_H, V_L) = \frac{M(l, V_H + V_L)}{V_H + V_L}$ is the probability that a vacancy will be initially matched with a worker. Note that the firm's choice of v does not affect $h(V_H, V_L) E[\pi_f | V_H, V_L]$ under the assumption that each firm is atomistic. The first order condition to the vacancy problem equates the marginal cost of the vacancy to the expected equilibrium profits:

$$c'(v) = hE\left[\pi_f | V_H, V_L\right]$$

so the equilibrium share of vacancies from high learning firms is:

$$s = \frac{v (hE [\pi_H | V_H, V_L])}{v (hE [\pi_H | V_H, V_L]) + v (hE [\pi_L | V_H, V_L])}$$

where $v(\pi)$ is the vacancy supply function given (expected) profits π . s is always increasing in ψ if and only if

$$\frac{v'(\lambda\psi)\,\lambda}{v(\lambda\psi)} > \frac{v'(\psi)}{v(\psi)}$$

for all $\lambda > 1$. A sufficient (but not necessary) condition for this is if v is log-convex, which is satisfied if $v = e^{\pi}$ (i.e., c' is log), or if $v(\pi) = \pi^{-a}$ for some a > 0.

Define firm's solution to the optimal vacancy problem as $v_f^*(V_H, V_L)$, where (V_H, V_L) are the total number of vacancies posted by high learning and low learning firms in the market. The equilibrium is the fixed point where $v_f^*(v_H, v_L) = v_f \ \forall f \in \{H, L\}$. To see that this fixed point exists under my assumptions, I can rewrite the maximization problem as

$$\max_{v} h\left(v_{H}+v_{L}\right) E\left[\pi\left(v_{H},v_{L}\right)\right] v-c\left(v\right),$$

noting that the market level of vacancies imposes two negative externalities for the firm. First, excess vacancies increase the match rate on the secondary market, which drives up wage competition for incumbent workers. Second, excess vacancies lower the initial fill rate for new workers. There is also an offsetting force, where the share of firms that are high learning determine the degree of adverse selection on the secondary market. Expected profits are clearly decreasing in the total number of vacancies due to the first two mechanisms, so it suffices to ensure that the net effect of adding high learning employer vacancies does not induce unraveling (e.g., the adverse selection effect is not so strong as to push the slope of the best response curve for high learning vacancies above 1).

To be precise about this condition, notice that clearly $h(v_H + v_L)$ is decreasing in the market v_H . So, a sufficient (but not necessary) condition for the best response function to be decreasing in v_H is if

$$\frac{\partial E\left[\pi\left(v_H, v_L\right)\right]}{\partial v_H} \le 0.$$

Since the expected profits for each filled vacancy is the weighted average of the profits from a worker of known high ability and the profits from the average worker (and neither α or the firm-level Qdepends on the market characteristics), it suffices to consider whether

$$\frac{\partial \pi_1}{\partial v_H}, \frac{\partial \pi_2}{\partial v_H} \le 0 \iff \frac{\partial w_1}{\partial v_H}, \frac{\partial w_2}{\partial v_H} \ge 0.$$

Secondary market profits are zero in any equilibrium, so the conditions that ensure firm profits are sufficiently well behaved are also exactly the conditions that ensure offered wages are increasing. Now,

$$\frac{\partial w_1}{\partial v_H} = \frac{\partial g}{\partial v_H} E_{\alpha'} \left[f_1(\theta) \right] + g \frac{\partial E \left[f_1(\theta) |\alpha' \right]}{\partial v_H}$$

$$\frac{\partial w_2}{\partial v_H} = \frac{\partial g}{\partial v_H} \left[\kappa f_2(\theta_H) + (1-\kappa) E \left[f_1(\theta) |\alpha' \right] \right] + g \left(1-\kappa\right) \frac{\partial E \left[f_1(\theta) |\alpha' \right]}{\partial v_H}.$$

Observe that the wages of promoted workers will be more insulated from adverse selection in secondary market than the wages of unknown workers since their expected outside option includes the possibility of obtaining their true product, so a sufficient condition for both derivatives to be positive is for $\frac{\partial w_1}{\partial v_H} \ge 0$. A simple rearrangement of the derivative then yields the assumed condition in the proposition:

$$-\frac{\partial E\left[f_{1}\left(\theta\right)|\alpha'\right]}{\partial v_{H}}/E\left[f_{1}\left(\theta\right)|\alpha'\right] \leq \frac{\partial g}{\partial v_{H}}/g.$$

Conditional on the equilibrium existing, it's straightforward to show that increasing the productivity term ψ increases the best response to any market-level vacancy, so the equilibrium number of vacancies will increase as well.

Similarly, the assumptions on secondary market matching are sufficient to ensure that the probability of receiving a secondary market offer, g, is increasing in the number of initial vacancies. So clearly, workers' likelihood of re-employment is higher in more productive regions. Meanwhile, incumbent wages increasing with high learning employer vacancies is exactly the sufficient condition that guarantees $\partial E[\pi(v_H, v_L)]/\partial v_H \leq 0$, so it follows that the same conditions that ensure the existence of the vacancy creation problem also ensures that occupational wages at the incumbent employers are weakly increasing as well.

C Additional Empirical Details

C.1 Data Construction

C.1.1 Annual worker panel

My primary data on worker earnings, job characteristics, and employer characteristics come from the universe of formal employment contracts in the RAIS data. Each observation in the raw data is a single employment contract within a state and year, so my first step is to construct an annual panel of workers' employment histories. Each observation in the annual panel is a worker's primary employment contract for that year, and the resulting dataset serves as the basis for all subsamples and derived measures in my project.

To construct the unique worker-by-year panel of primary employment, I consider all employment contracts that covered at least six months over the year, entailed at least 20 contracted hours of work per week, and paid non-zero earnings. In the cases when there are multiple recorded employment contracts for the same worker and year that satisfy these selection criteria, I choose the employment contract covering the longest duration (in months), and I break any subsequent ties by selecting the contract with the highest average monthly earnings.

My preferred earnings measure is the average nominal monthly earnings over the employment contract. Where defined, this measure is highly correlated with the December monthly earnings measure that has been used in the literature (the correlation coefficient between these measures in logs is above .97), but average monthly earnings the additional advantage of being defined for partial employment spells in a year that ended before December.

I classify any years when a worker is not in the annual panel as years when the worker is out of the formal labor market. Correspondingly, I consider the worker to not be working as a supervisor in those years. I impute worker characteristics in years when the worker is not in the annual panel by using the last known observation (for gender, education, state, race, and birth year). Meanwhile, I impute a worker's counterfactual earnings by annually compounding the worker's last known earnings by the average wage growth for the worker's last known state-by-baseline-occupationgroup. Finally, I assign a reason for separation to the out-of-formal-labor-market spell using the reason for separation field of the most recent employment contract. Employer-initiated separations are employer terminations with or without just cause (excluding contract expirations). Workerinitiated separations are voluntary worker separations. I combine separations for any other known reason (including contract expirations) into a single category.

C.1.2 Classifying promotions

I use the CBO-02 occupation codes recorded in the RAIS for each employment contract from 2003 onwards to define promotions. The occupation codes follow a consistent hierarchical structure, so I can define supervisory jobs from the structure of the occupation codes themselves. As a check, I can also define supervisor jobs from the text of the job titles (by finding all job titles that contain

the term *supervisores*). The results from these two classification methods are identical, which is reassuring about the consistency of the occupation classification system.

The CBO-02 system classifies all occupations into a 6-digit occupational code. The first two digits of the occupation code indicate the main occupation group, which are generally broad classes of jobs like metalworkers, textile workers, or public services workers. Within production-level occupational groups, a third digit of "0" in the occupation code is reserved for the supervisors in the occupational group, whereas all other values refer to other sub-groups within the occupation that do not necessarily have a clear vertical interpretation relative to each other. Individual occupations are further differentiated by the three additional digits following these three base digits.

Within production-level occupation groups (CBO-02 codes starting with 41-99), all but four occupational groups contain supervisor occupations. On the other hand, none of the civil, managerial, professional, or technical-level occupational groups (CBO-02 codes starting with 01-39) contain supervisor occupations. So, occupational groups with observable lines of progression can be considered to be a proper (but nearly complete) subset of the production worker-level (*trabalhadores*) occupation groups.

C.1.3 Estimating promotion propensity

As discussed in Section 4.1, I classify firms as high or low promotion firms based on their compositionadjusted promotion rates between 2004 and 2006. Specifically, for each of the three years t, I consider all workers from the annual panel between the ages of 25 and 50 who were in formal employment at a non-public sector firm in years t - 1 and t. I further restrict the sample to all workers who remained in the same broad occupation group in both years, which bolsters the interpretation that these promotions reflect vertical job changes. I estimate η_{jt} on the sample using Equation 5 as the residual promotion rate for firm j at time t after adjusting for worker characteristics and differences in promotion rates in different occupation groups. Finally, to minimize measurement error or the contribution of year-specific shocks, I restrict the set of firms to those that had at least 10 workers in the estimation sample for each of the three years, and I define $\eta_j = E[\eta_{jt}]$ as the firm's average promotion residual over those three years.

It's worth noting two additional details implicit in the baseline approach. I do not restrict the sample to workers who are at the same firm in both years, so some promotions in the data are external promotions where a worker was working as a line worker in a firm in year t - 1 and as a supervisor in a different firm in year t. Furthermore, in the case of these external promotions, I attribute the promotion to the firm where the worker received the promotion (i.e., the firm in year t), which is consistent with my interpretation and model. However, although the inclusion of external promotions slightly increases power, these two details are inconsequential for my results. Table A.1 compares the correlation in firms' promotion rates across a variety of alternate classification methods. The measures are highly correlated with each other. As an additional check, Figure A.18 compares the treatment effect estimates when I classify firms based on their internal promotion rates - the effects on promotions and turnover are slightly attenuated but otherwise similar.

C.1.4 Measuring mass layoffs

I follow the literature on using linked employer-employee data to identify mass-layoff events. I first compile a firm-level panel of employment counts by aggregating the worker-level annual panel to the firm-year level, and I identify large employment drops or firm closures using the criteria from Schmieder et al. (2020). Mass-layoff events are when a firm with at least 50 employees experience at least a 30% drop in employment or disappear from the data altogether in the following year. Firm identifiers are not always longitudinally consistent, so reorganizations or spinoffs may be mistakenly classified as layoff events. I follow the literature to exclude these alternate scenarios by dropping any layoff event where at least 20% of displaced workers go to the same firm in the following year.

C.1.5 Constructing the local hiring IV

I construct a panel of total employment and new hiring at the firm-by-municipality level by aggregating the worker-level annual panel. Since employment contracts specify both the firm and location of the employment establishment, the mapping from workers to the firm-municipality is clear. New hires are defined as the total number of workers who are working in the firm-municipality and were working at a different firm in the year prior. So, this measure excludes within firm transfers across municipalities, as well as any brief employment spells that would not be classified as the worker's primary employment for the year. To ensure that the hiring shares are informative, I restrict my attention to municipalities that have at least 1000 workers and at least 200 new hires each year.

For a worker i in mover cohort c and municipality m, I calculate their jack-knife local hiring share instrument as

$$z_{mc} = \frac{G_{mc}^{H} - \sum_{J(i')=J(i)} H_{i'}}{G_{mc}^{H} + G_{mc}^{L} - \sum_{J(i')=J(i)} 1}.$$

where G_{mc}^{H} , G_{mc}^{L} are the total new hires by high and low promotion firms, respectively, and i' are other movers in the same cohort. Since workers in my analysis sample are included in the new hire totals, I avoid the reflection problem that would arise from this functional dependence by excluding *all* new hires from worker *i*'s destination firm from the numerator and denominator. As a result, the interpretation of z_{mc} is the local hiring share by high promotion firms *excluding* the worker's destination firm. Technically, z_{mc} varies by destination firm due to the jack-knife procedure. However, this variation is minor, so I slightly abuse notation to focus on the main source of variation.

C.2 Calculating Average Treatment Effects

I combine workers from multiple cohorts to increase the precision of my estimates and to ensure that I am capturing an average treatment effect that is representative across cohorts. However, researchers have cautioned that pooling treatment effects in designs with staggered treatment timing may yield unintuitive and potentially negative weighting of the underlying treatment effects.

To address these concerns and make the relevant comparisons clear, I allow all coefficients in the estimation equation to vary arbitrarily with the worker's cohort. This clearly emphasizes that all identification of treatment effects over time come solely from comparisons between workers who are in the same cohort (e.g., I compare workers who moved to a high promotion firm in 2009 to workers who moved to a low promotion firm in 2009). Furthermore, I can combine estimated treatment effects across cohorts using explicitly specified weights to estimate an average treatment effect across all mover cohorts.

I define my estimand of interest as the average effect across all cohorts. Correspondingly, the

combined estimate of the cohort specific treatment effect $\beta_{c\tau}$ at event time τ is

$$\beta_{\tau} = \left[\sum_{c=2008}^{2012} \beta_{c\tau}\right] / 5.$$

In the data, the number of workers in each cohort varies slightly, so an alternative is to weigh each cohort's treatment effect by the number of workers in the cohort. However, using uniform weighting across cohort years is more straightforward, and ensures that any differences in estimates across subgroups are not driven by any differences in the composition of workers across cohorts.

In practice, the difference between all of the possible approaches is small. Estimating Equation 6 by pooling the treatment effect coefficients yield similar estimates. This is due to two reasons. First, cohort-specific treatment effects are already reasonably similar. In addition, the share of workers moving to high versus low promotion firms each year is also stable, so the OLS weights are roughly comparable to the uniform weights.

C.3 Structural Quantification

C.3.1 Estimation method

The model yields nonlinear expressions for the means and treatment effects of high learning firms on promotions and turnover. I match these expressions to their empirical analogs using classical minimum distance, which requires numerically optimizing the nonlinear objective function in Equation 12. The objective function is straightforward to compute given that the model expressions are in closed form, yet numerical optimization can run the risk of hitting local rather than the global minima of the objective function.

I use the following algorithm to calculate my model parameters. First, I draw a starting guess for the parameters $\theta = (\bar{Q}_H, \bar{Q}_L, \alpha, g)$ randomly from a uniform distribution. I then numerically minimize the objective function using the Nelder Mead algorithm from the R package nloptr with a stopping criterion for the relative change of 10^{-12} and a constraint that each parameter lies on the interior between 0 and 1. I repeat the process 100 times, drawing a new random starting value each time, and I select the solution with the smallest objective across all the starting value draws.

The nonlinear model performs reasonably well in my setting. Across the 100 different starting values, the numerical algorithm reaches an objective below 10^{-5} in 68 cases. The maximum standard deviation for any parameter estimate across these 68 cases is approximately 10^{-13} and the mean objective is approximately 10^{-22} .

C.3.2 Sensitivity of parameter estimates to moments (Λ)

Although the estimation of the parameters requires optimizing a nonlinear function, the choice of the minimum distance estimator ensures that all identification for the model parameters ultimately comes from the four empirical moments (conditional on the calibrated parameters). To help increase the transparency of the model estimates, I report the *sensitivity* of the model parameters to the matched moments, as defined by Andrews et al. (2017), in Table A.6.

Formally, the sensitivity measure in the classical minimum distance estimator is

$$\Lambda = \left(G'WG\right)^{-1}G'W,$$

where W is the chosen weighting matrix and G is the Jacobian of the model equations $h(\theta)$ from Equation 12. The interpretation is that for a local perturbation of the empirical moments that converges to η , the first-order asymptotic bias in the estimated parameters is

$$E\left[\tilde{\theta}\right] = \Lambda \eta.$$

For more details, see Proposition 2 of Andrews et al. (2017). As expected, the estimated parameters are most sensitive to the treatment effect estimates of high promotion firms, particularly their effects on promotions.