IRLE
Institute for Research on Labor and Employment

IRLE WORKING PAPER
#105-21
August 2021

The Dynamics of Referral Hiring and Racial Inequality: Evidence from Brazil

Conrad Miller and Ian Schmutte

Cite as: Conrad Miller and Ian Schmutte. (2021). "The Dynamics of Referral Hiring and Racial Inequality". IRLE Working Paper No. 105-21.
http://irle.berkeley.edu/files/2021/09/The-Dynamics-of-Referral-Hiring-and-Racial-Inequality-Evidence-from-Brazil.pdf

The Dynamics of Referral Hiring and Racial Inequality: Evidence from Brazil

Conrad Miller[*]

Ian Schmutte[†]

August 22nd, 2021

Abstract

We study how referral hiring contributes to racial inequality in firm-level labor demand over the firm's life cycle using data from Brazil. We consider a search model where referral networks are segregated, firms are more informed about the match quality of referred candidates, and some referrals are made by non-referred employees. Consistent with the model, we find that firms are more likely to hire candidates and less likely to dismiss employees of the same race as the founder, but these differences diminish as firms' cumulative hires increase. Referral hiring helps to explain racial differences in dismissals, seniority, and employer size.

[*]Haas School of Business, University of California, Berkeley and NBER (email: ccmiller@berkeley.edu)

[†]University of Georgia (email: schmutte@uga.edu). We thank Joaquín Fuenzalida for excellent research assistance. We thank Kathryn Anderson, Sydnee Caldwell, Jonathan Leonard, David Levine, Ricardo Perez-Truglia, Isaac Sorkin, Guo Xu, and seminar participants at the 2020 Southern Economic Association meetings, the OZ Virtual Seminar, Notre Dame, Cornell, UC Berkeley, Claremont McKenna College, Rutgers, Indiana, LSE, West Point, Chicago Harris, UCSB, the Southeastern Micro Labor Workshop, Inter-American Development Bank, Yale, and NBER Personnel for comments. Miller gratefully acknowledges financial support from the Upjohn Institute and the UC Berkeley Institute for Research on Labor and Employment.

1 Introduction

Social networks affect which and when job seekers find work, the types of jobs they end up in, and the wages they are paid. They channel information to both employers and workers about the existence of potential matches and the quality of those matches. There are also persistent racial differences in the use and productivity of referrals—differences that arise because social networks tend to form along existing lines of social and economic stratification (McPherson et al., 2001; Pedulla and Pager, 2019). Theory suggests that the widespread practice of referral hiring may perpetuate or exacerbate racial disparities in labor market outcomes by funneling opportunities within groups (Calvo-Armengol and Jackson, 2004; Bolte et al., 2020; Okafor, 2020). Yet the extent to which referral hiring contributes to racial inequality in practice is not well understood. For example, while several single-firm case studies find that the racial composition of a firm's referral hires tends to reflect the racial composition of the firm's incumbent employees (Fernandez et al., 2000; Fernandez and Sosa, 2005; Petersen et al., 2000), it is not clear to what extent referral hiring shapes the composition of employees in the first place or persistently favors specific groups over time.

We study the firm-level implications of referral hiring for racial inequality in labor demand and match quality using detailed employer-employee data from Brazil, a country with well-documented racial disparities in employment rates and wages.[1] We emphasize how the role of referral hiring evolves over a firm's life cycle. We present a simple job search model where (a) social networks are racially segregated, (b) firms are more informed about the match quality of job seekers who are referred by an incumbent employee, and (c) at least some referrals are made by non-referred employees. We confirm four key predictions of the model. First, firms with white founders are more likely to hire white employees than comparable firms with nonwhite founders. Second, these differences disappear as firms' cumulative number of hires increases. Third, firms are less likely to dismiss recent hires of the same race as the firm's founder, indicating those hires are less likely to be an ex-post poor match. Fourth, racial differences in dismissal rates are also decreasing in a firm's cumulative number of hires. Yet few firms hire enough employees to reach convergence in their racial composition of hires or dismissal rates. We then show that our findings, given that founders are disproportionately white, help to explain three stylized facts about racial differences in labor market outcomes: nonwhite workers are more likely to be dismissed by their employers, have less seniority, and sort to larger employers than white workers.

We first describe a simple job search model based on Morgan and Várdy (2009) where a firm posts job vacancies and is matched with job seekers either through referral or the external market. At the firm's entry, referrals are drawn from the founder's network. Once the firm has hired additional employees, referrals are drawn from other incumbent employees as well. We assume that referral networks are segregated so that the racial composition of referral candidates tends

[1] See, for example, Silva (1980, 1985); Lovell (1994); Cavalieri and Fernandes (1998); Arcard and d'Hombres (2004); Arias et al. (2004); Matos and Machado (2006); Garia et al. (2009); Reis and Crespo (2015); Gerard et al. (forthcoming); Derenoncourt et al. (2021).

to reflect the racial composition of incumbent employees. Following much of the literature (Topa, 2019), we also assume that a job seeker's match-specific productivity is more uncertain when they are matched to the firm via the external market. To match the observation that large employers invest more in formal hiring methods and find a smaller share of their hires via referral (Barron et al., 1987; Holzer, 1987a; Marsden, 1994; Rebien et al., 2020), we further suppose that the firm can invest in a (fixed cost) recruitment and screening technology that reduces the relative uncertainty associated with external market candidates.

The model has stark predictions for how a firm's racial composition of hires and racial differences in dismissal rates evolve over time. The first prediction is that firms with white founders are more likely to hire white employees than comparable firms with nonwhite founders. This follows immediately from the assumption that referral networks are racially segregated. The second prediction is that the racial composition of hires for firms with white and nonwhite founders converges as firms' cumulative number of hires increases. The correlation between the race of a firm's founder and the firm's racial composition of recent hires weakens with the firm's cumulative number of hires for two reasons. First, employees hired via the external market eventually provide referral candidates themselves, pushing the composition of hires toward the composition of external market candidates. Second, firms that expect to hire more employees invest in hiring technology that reduces the referral share of hires.

The third prediction is that firms are less likely to dismiss recent hires of the same race as the firm's founder. In the model, referral hires are less likely to be poor matches ex-post and hence should have lower turnover rates. This turnover advantage should diminish with job tenure as both the employer and worker learn about match quality. At firms with white founders, white hires are more likely to be referrals than nonwhite hires, while the opposite is true at firms with nonwhite founders. The fourth prediction is that racial differences in dismissal rates are diminishing in a firm's cumulative number of hires. The referral share of recent white and nonwhite hires will become more similar as employers' cumulative hires increase, both because the racial composition of referral and non-referral candidates become more similar and because the overall referral share of hires is smaller at firms that hire more workers.

We evaluate the model using linked employer-employee data from Brazil. First, we test two key assumptions: (a) referral networks are racially segregated, and (b) the use of referral hiring is decreasing in employer size. While both assumptions have been shown to hold in other settings, we show that they hold in our data as well. The idea behind our tests is that if referrals are an important hiring channel, we should observe that firms are more likely to hire job seekers with a social connection to one of their incumbent employees. Since we do not observe referrals directly, we proxy for social connections by identifying pairs of workers that worked together in the past.[2]

[2] A growing body of research establishes that job referral effects can be credibly measured using models that proxy for social connections using information on past coworkers (Cingano and Rosolia, 2012; Hensvik and Skans, 2016; Saygin et al., 2021; Glitz, 2017), residential neighbors (Bayer et al., 2008; Hellerstein et al., 2014; Schmutte, 2015), and family ties (Kramarz and Skans, 2014). We focus on past coworking connections because the RAIS data do not allow us to track residential location or family relationships.

Following Eliason et al. (2020), we measure the effect of social connections on where workers sort by comparing the destinations of workers that separate from the same employer but have different social connections. We refine their approach by comparing outcomes when a true coworker is present to the outcomes for what we call *placebo* coworkers—those pairs that were previously employed in the same job, but at slightly different times. Consistent with racially segregated referral networks, we find the effects of social connections on hiring probabilities are substantially larger when the potential hire and connected incumbent are of the same race.[3] We also show that the effects of social connections are decreasing in destination plant size.

Second, we show that the racial composition of hires for firms with white and nonwhite founders differs substantially at entry but converges as their cumulative number of hires increases. For the same local labor market and occupation, early hires at firms with white founders are about 40% more likely to be white than early hires in similar positions at firms with nonwhite founders. Yet at firms with a white founder, the nonwhite share of hires increases sharply with cumulative hires, and the nonwhite share of hires *decreases* with cumulative hires at firms with nonwhite founders. These patterns hold within a balanced panel of firms. The racial composition of hires at firms that have hired 350 or more workers after their year of entry is unrelated to founder race. Yet most firms do not reach this scale—five years after entry, fewer than 1% of remaining firms have made this many hires, accounting for 14% of hires.

Third, we confirm model predictions for turnover patterns. Consistent with firms having superior information about the match quality of referral job candidates, recent hires with preexisting social connections at their employer are dismissed at substantially lower rates than recent hires without those connections, and recent hires with placebo connections in particular. This turnover advantage decreases over the job spell. We also find that nonwhite employees are dismissed at higher rates than white employees at firms with white founders, while the opposite is true for firms with nonwhite founders. These racial differences in dismissal rates are declining in a firm's cumulative number of hires and over the course of the job spell.

We then examine the implications of our findings for racial inequality in the labor market. A key market feature for understanding the aggregate implications of referral hiring is that racial disparities in labor market outcomes often coincide with racial differences in entrepreneurship rates. In Brazil, white men and women engage in entrepreneurship at rates about twice as high as nonwhite men and women, where we define entrepreneurship as running a formal or informal business with at least one paid employee. This suggests that small or young firms will disproportionately favor white job seekers in hiring.

Following this reasoning, we discuss three stylized facts about racial differences in labor market outcomes that are consistent with the dynamics of referral-based hiring: relative to white workers, nonwhite workers (1) are more likely to be dismissed by their employers, (2) have less seniority, and (3) sort to larger employers. All three patterns are driven by employers with white founders.

[3]These results are also related to a literature showing that close connections are more valuable for job-finding (Gee et al., 2017a).

Interestingly, the first and third patterns also hold for black and white workers in the United States (Lang and Lehmann, 2012; Cavounidis et al., 2021; Holzer, 1998; Miller, 2017), suggesting referral hiring may have similar dynamic implications for racial inequality there.[4]

There are alternative explanations for some of our findings, but we show that none can easily match the combination of findings we document. For example, discriminatory founder or coworker preferences can potentially explain why firms with white founders have more white employees than comparable firms with nonwhite founders, and why larger firms are more racially diverse (if less discriminatory firms are more productive). But it is unclear how that mechanism would generate racial differences in dismissal rates or convergence between firms with white and nonwhite founders over time in hiring and dismissal behavior. Moreover, mobility patterns suggest that white and nonwhite workers do not have systematically different preferences over employers, at least as characterized by founder race and cumulative hires (Bagger and Lentz, 2018; Sorkin, 2018). There is one alternative mechanism that could yield similar predictions to the referral mechanism emphasized here: hiring managers are better at screening same-race applicants (Giuliano et al., 2011; Åslund et al., 2014; Benson et al., 2019, see also Fisman et al., 2017). We view these mechanisms as similar in that they both posit that social (or cultural) proximity affects screening ability. We focus on referral hiring because aspects of referral hiring that are essential to the model examined here—that referral networks are racially segregated and employers have additional information about the match quality of referral candidates—are empirically supported in our setting and the broader literature. By contrast, the premise that managers are better at screening same-race applicants is more speculative.

Our findings have implications for policy. First, they provide a novel rationale for affirmative action policies. As firms mature, the racial composition of their hires (slowly) converges to the composition of the external market. A policy that incentivizes firms to hire workers from groups underrepresented at the firm relative to the external market would accelerate this process. Moreover, temporary affirmative action policies will have persistent effects as in Miller (2017). Second, our findings suggest that market frictions that affect the size distribution of firms will have implications for racial inequality in the labor market. For example, if small, productive firms are unable to expand to their efficient size due to some resource misallocation (Restuccia and Rogerson, 2017), these firms are also less likely to have a racially diverse workforce. The logic of our framework suggests that the aggregate costs of misallocation will be disproportionately borne by groups underrepresented among entrepreneurs.

We add to a growing economics literature on the role of referral networks. We contribute methodologically by using placebo coworker connections to identify referral effects in hiring. Our design was inspired by several previous papers that use similar approaches. Hensvik and Skans (2016) compare true and placebo coworkers to infer the characteristics of workers that receive and provide referrals, and the effects of being referred on job outcomes. Caldwell and Harmon (2019)

[4] To the best of our knowledge, seniority or relative tenure within the same employer by race has not been studied in the United States.

study the effect of coworker networks on job mobility and earnings and compare outcomes based how long ago the coworker relationship took place. San (2021) compares hiring outcomes when a social connection is present and could potentially provide a referral relative to periods just after they have left the firm. Relative to these papers, our focus on comparing past true and placebo coworking relationships to identify hiring effects is novel. Furthermore, using the monthly detail in the data, we show a sharp discontinuity in hiring for workers who actually overlap to those who almost, but did not, work together. These details add broad support and credibility to the literature following Bayer et al. (2008) that uses variation in the amount of social distance between workers to infer social interaction effects in hiring.

We also contribute to an extensive literature on the role of referral networks in driving persistent between-group inequality.[5] Topa (2001) shows that the concentration of unemployment across neighborhoods in Chicago can be explained by a model where there are neighborhood interactions in job search, consistent with the model of Calvo-Armengol and Jackson (2004). A subset of this literature focuses on employer behavior and how employers' reliance on referrals influences inequality. These papers typically study applications to a single firm (Fernandez et al., 2000; Fernandez and Sosa, 2005; Petersen et al., 2000). We argue that the racial composition of a firm's referral hires depends critically on the firm's initial conditions (which we proxy with the race of the founder) but converges to that of the external market over the course of the firm's life cycle.[6]

This paper provides an alternative rationale for models of statistical discrimination that assume decision-makers can screen one group of candidates with more precision than another group (Aigner and Cain, 1977; Lundberg and Startz, 1983; Cornell and Welch, 1996; Morgan and Várdy, 2009). Prior work has justified this assumption by appealing to between-group differences in culture or communication. We offer a complementary explanation: small or young firms with white (nonwhite) founders screen white (nonwhite) job candidates with more precision on average because a higher share of their white (nonwhite) candidates are referrals.

Finally, we contribute to a literature on entrepreneurship and co-racial or co-ethnic hiring. Bates (2006) and Boston (2006) document that, in the United States, black business owners employ black workers at higher rates than white business owners, even within the same local labor markets. Both authors argue that increasing black entrepreneurship rates will reduce black unemployment rates. In contemporaneous work, Dias and Rocha (2021) document a similar pattern in Brazil and find that racial wage disparities are smaller in firms with nonwhite ownership. Kerr and Kerr (2021) study co-ethnic hiring by immigrant entrepreneurs in the United States. They find that the average new firm with five or more workers has a co-ethnic share of about 22.5%, with substantial variation by the entrepreneur's country of origin. Interestingly, they find similar co-ethnic employee shares

[5]While this literature has primarily focused on the United States, our analysis of Brazilian data is relevant. The key facts that motivate our model are based on studies from the United States, and there is some evidence that referral use in Brazil is very similar to the United States and other OECD countries (Gee et al., 2017b).

[6]Another subset of this literature focuses on job search behavior. A central question in this literature is whether black job seekers face lower returns to using network-based search methods and, if so, why (Holzer, 1987b; Fernandez and Fernandez-Mateo, 2006; Pedulla and Pager, 2019; DiTomaso, 2013). Our findings also suggest that nonwhite workers are more likely to be socially connected to large firms, which are less dependent on referral hiring.

several years after a firm's birth. By contrast, we focus on how the composition of hires evolves with an employer's cumulative hires to date.

The remainder of the paper is organized as follows. Section 2 sets up our job search model and derives predictions. In Section 3 we describe the Brazilian context and employer-employee data that form the basis of our study. In Section 4 we test predictions of the model. In Section 5 we address alternative interpretations of our findings. We examine the implications of our findings for racial inequality in Section 6. Section 7 concludes.

2 A Job Search Model with Referral Hiring

In this section we describe a simple job search model where a firm's incumbent employees provide otherwise unobservable information about the productivity of their social connections to the firm. Much of the structure closely follows Morgan and Várdy (2009). We derive predictions for how the racial composition of a firm's hires varies (1) with the race of the founder, (2) over time, and (3) with the firm's size. We also derive predictions for how racial differences in dismissal rates among recent hires vary with these factors.

We consider the sequential hiring decisions of a single firm, which must fill n vacancies. The wage for each position is fixed. Consider the firm's i^{th} vacancy. To fill a vacancy, the firm interviews randomly drawn candidates at a cost $k \geq 0$ per interview. With probability ω, the candidate was referred by a random member of the firm's existing workforce, and with probability $1-\omega$, the candidate applied through the external market. Let γ denote the pool a candidate is drawn from, with $\gamma = R$ for referred candidates and $\gamma = E$ for candidates drawn from external market.

Founder or candidate race is denoted by $\rho \in \{W, N\}$. Let $r_i^h \in \{W, N\}$ denote the race of the i^{th} hired candidate who is hired for vacancy $v(i)$. Let $r_j^e \in \{W, N\}$ denote the race of the candidate who *fills* vacancy j, meaning the candidate is both hired for vacancy j and retained after the probationary period, which is discussed below. The race of the firm's founder is given by $r_0^e \in \{W, N\}$. Let π_j denote the share of the incumbent workforce with race N ("nonwhite share") when the firm is filling vacancy j, where

$$\pi_j = \frac{1}{j} \sum_{\ell=0}^{j-1} \mathbb{1}_{\{r_\ell^e = N\}}.$$

We assume that incumbent employees only refer workers of the same race. Hence, the probability that a referral candidate for vacancy j is nonwhite is given by π_j. This assumption is stark but is consistent with well-documented racial homophily in referral networks and social networks more broadly (McPherson et al., 2001; Fernandez and Fernandez-Mateo, 2006; Hellerstein et al., 2011, 2014; Brown et al., 2016).[7] Let $\bar{\pi}$ denote the nonwhite share of external market candidates.

A candidate's match-specific productivity, θ, equals one if the candidate can perform the job

[7] We would reach similar conclusions as long as the probability that a referral candidate is nonwhite is increasing in π_j.

and zero if the candidate cannot. Let $p \equiv Pr(\theta = 1)$ denote the probability that the candidate drawn randomly from the population of job seekers can perform the job. We assume this probability is independent of the pool a candidate is drawn from (γ) and candidate race (ρ).

At the interview stage, the firm receives a noisy signal S_γ for the candidate's productivity, $S_\gamma = \theta + \epsilon_\gamma$, where ϵ_γ is normally distributed with mean zero and variance σ_γ^2. Following the literature, we assume that this signal is more precise for referral matches so that $\sigma_R^2 < \sigma_E^2$ (Topa, 2019). Later we will allow firms to improve the precision of external market candidates at some cost.

For each vacancy, the timing is as follows. In period 1, the firm draws a random candidate and conducts an interview. On the basis of the candidate's signal, s, and pool, γ, the firm forms a posterior belief, q, about the candidate's match-specific productivity. The firm then decides whether to hire the candidate, and period 1 ends.

In period 2 and all subsequent periods, if the firm did not hire in the previous period, the firm interviews a new candidate and the process proceeds as before. If the firm did hire in the previous period, the employee's productivity θ is revealed to the firm. If $\theta = 1$, the employee is retained forever and the firm moves on to fill the next vacancy, if there is one. In that case, the firm receives a payoff with net present value $v > 0$. If $\theta = 0$, then the firm receives a payoff with net present value $-w < 0$ if it retains the employee and incurs cost $c > 0$ if it dismisses the employee. Throughout, we assume that $c < w$ so that it is always optimal to dismiss unproductive employees. Finally, we assume that the employer is risk neutral and has a discount factor $\delta \in (0, 1)$.

As Morgan and Várdy (2009) show, the firm's optimal strategy is to impose a uniform success probability threshold, q^*, when deciding whether to hire a candidate. This threshold does not depend on the candidate's pool or race. Define $q_\gamma(s)$ as the firm's posterior belief that a candidate from pool γ with signal s can perform the job; that is, $q_\gamma(s) \equiv Pr(\theta = 1|S_\gamma = s)$. By Bayes' rule, this can be written as

$$q_\gamma(s) = \frac{\pi[(s-1)/\sigma_\gamma]p}{\phi[(s-1)/\sigma_\gamma]p + \phi[s/\sigma_\gamma](1-p)},$$

where $\phi(\cdot)$ denote the probability density of a standard normal random variable.

Let $s_\gamma(q)$ denote the signal realization that corresponds to a given success probability q. Before the realization of the signal, but after the firm observes a candidate's pool, the success probability $Q_\gamma = q_\gamma(S_\gamma)$ is a random variable. Now, let $G_\gamma(\cdot)$ denote the cumulative distribution function (CDF) of Q_γ. $G_\gamma(q)$ is given by

$$G_\gamma(q) = p\Phi\left(\frac{s_\gamma(q) - 1}{\sigma_\gamma}\right) + (1-p)\Phi\left(\frac{s_\gamma(q)}{\sigma_\gamma}\right),$$

where $\Phi(\cdot)$ denotes the CDF of a standard normal random variable.

Similarly, let $G(\cdot)$ denote the CDF of the success probability prior to observing a candidate's pool or signal, where $G(q) = \omega G_R(q) + (1-\omega)G_E(q)$.

2.1 Composition of Hires

We now consider the probability that a hire is nonwhite. Let α denote the probability that a hire is a referral. α is given by

$$\alpha = \frac{\omega(1 - G_R(q^*))}{\omega(1 - G_R(q^*)) + (1 - \omega)(1 - G_E(q^*))}.$$

Hence, the probability that hire i is nonwhite is given by

$$P(r_i^h = N) = \alpha\pi_{v(i)} + (1 - \alpha)\bar{\pi}.$$

Note that for an employer where $\pi_{v(i)} < \bar{\pi}$, the nonwhite share of hires is decreasing in the referral share of hires, α.

Similarly, the probability that a *successful* hire for vacancy i is nonwhite is equal to $\alpha'\pi_{v(i)} + (1 - \alpha')\bar{\pi}$, where α' is the probability that a *successful* hire is a referral. In steady state, the nonwhite share of incumbent employees ($\pi_{v(i)}$) is equal to the nonwhite share of the external market ($\bar{\pi}$).

Finally, we allow firms to adjust the precision of the productivity signals they receive in the referral market so that signal precision $h_E = \frac{1}{\sigma_E^2}$ is a function investment, c_p.[8] We assume that these costs are fixed relative to a firm's number of vacancies n. Hence, larger firms, or those with more vacancies, invest more in screening precision and have a lower value of h_E. While this assumption is not essential to generating the key predictions of the model, we include it to match the stylized fact that large employers use more formal recruiting and screening methods and find a smaller share of their hires via referral (Holzer, 1987a; Marsden, 1994; Rebien et al., 2020). Morgan and Várdy (2009) show that if the firm is sufficiently selective (meaning q^* is sufficiently large), then an increase in a group's signal precision will increase the group's share of hires.[9]

We test four predictions about the composition of hires that derive from this framework. Two predictions follow from our assumptions that the costs of improving signal precision for external market candidates is fixed (and hence large firms hire a smaller share of their workers via referral) and referral networks are segregated:

1. The share of hires made via referral is declining in employer size.

2. The nonwhite share of referral hires is increasing in the nonwhite share of incumbent employees.

The next two predictions relate to how an employer's racial composition of hires varies over time and with employer size, as given by n. In particular, the composition of an employer's hires moves closer to that of the external market as either cumulative hires or total vacancies increase:

[8] Galenianos (2013) and Miller (2017) also allow the firm to control the precision of signals for external market candidates at some cost.

[9] There are alternative potential reasons that large employers are less likely to hire via referral that we do not model here. Small employers may be more risk averse. The arrival rate of referral candidates may be declining in employer size if, for example, the referral networks of coworkers are increasingly redundant.

3. The nonwhite share of hires converges to $\bar{\pi}$ as cumulative hires increases.

4. The nonwhite share of hires converges to $\bar{\pi}$ as employer size increases.

For employers with *white* founders, the expected nonwhite share of hires is *increasing* in cumulative hires and employer size.[10] For employers with *nonwhite* founders, the expected nonwhite share of hires is *decreasing* in cumulative hires and employer size.

2.2 Dismissal Rates

An additional four predictions follow from our assumption that the firm can screen referral candidates with more precision than external market candidates. Morgan and Várdy (2009) show that the group with lower screening precision is dismissed at higher rates. This prediction is standard in the literature (Brown et al., 2016; Topa, 2019). Moreover, this difference in dismissal rates diminishes over the job spell as the firm learns about employee productivity. This prediction is particularly stark in our model, where all employees with $\theta = 1$ are retained forever. However, the same prediction holds in more general models where productivity is continuous or employee productivity is revealed more gradually.

5. Referral hires have lower turnover rates than external market hires.

6. The referral turnover advantage is decreasing in job spell tenure.

7. Within-employer racial differences in dismissal rates are declining in cumulative hires and employer size.

8. Conditional on cumulative hires and employer size, racial differences in dismissal rates are decreasing in job spell tenure.

For employers with *white* founders, expected dismissal rates are higher for nonwhite hires than white hires. That is because white hires are more likely to be hired via referral, while the opposite is true for employers with nonwhite founders.

3 Context and Data

Like the U.S., Brazil's labor market exhibits significant racial disparities in wages and segregation in employment (Hirata and Soares, 2020; Gerard et al., forthcoming). However, Brazil has few regulations that protect workers against employment discrimination in the private sector on the

[10] The prediction for employer size is consistent with the literature that examines the effects of formal screening devices on hiring outcomes, including the racial composition of hires. Autor and Scarborough (2008) show that the introduction of job testing at a large retail firm did not reduce minority hiring despite minorities performing significantly worse on the test and generated productivity gains for both minority and non-minority hires. Holzer et al. (2006) and Wozniak (2015) argue that the use of criminal background checks and drug tests increases black hiring by providing information that is perceived to be more relevant for black candidates.

basis of race (Machado et al., 2019). Therefore, the differences we show in hiring patterns by race are unlikely to be shaped by regulatory pressure and instead reflect market or social institutions. We conduct our analysis using administrative linked employer-employee data from Brazil: the *Relação Anual de Informações Sociais* (RAIS), which include a remarkable amount of detail on the characteristics of both workers and their employment contracts.

3.1 Legal and Social Context

Brazil was founded as a race-based slave society and has persistent racial disparities across many socio-economic outcomes. For many decades after the end of slavery, Brazil maintained a national myth that it was a "racial democracy" in which racial disparities were incidental and transitory (Fiola, 1990). Brazil did not construct explicitly racist legal institutions equivalent to the Jim Crow era in the U.S., did not prohibit racial intermarriage, and did not operate under a genetic theory of racial superiority (Daniel, 2010). Perhaps as a result, the government has not adopted systematic affirmative action or equal opportunity policies that apply to the private sector.[11]

Given this history, it is not surprising that the sociology of race is also very different in Brazil than the United States. In Brazil, race is associated with skin tone and not so much a categorical trait fixed through inheritance. As a result, there is much more ambiguity and subjectivity in racial classification, which affects how race is measured in survey and administrative data. In official statistics, and in both of our main data sources, there are five main racial categories: *branco* (white), *preto* (black), *pardo* (brown), *amarelo* (yellow), and *indigena* (indigenous). However, the main axis of racial disparity is between the *branco* and the *preto* and *pardo* populations, who combined make up about 99% of the population. Therefore, like Cornwell et al. (2017), Hirata and Soares (2020), and Gerard et al. (forthcoming), we follow Telles (2004) in combining *pardo* and *preto* into a single "nonwhite" category and focus on comparing outcomes for white and nonwhite workers.

Brazil's labor markets are highly regulated in ways that affect our analysis. Workers on regular contracts have constitutionally guaranteed employment protection that kicks in after a 90-day probationary period. If they terminate a worker after the 90 days pass, firms must pay a fine proportional to a workers' completed tenure. Before that, they can fire workers at will. The presence of these termination costs affects firm's hiring decisions and the manner by which they evaluate workers during the probationary period. Arnold and Bernstein (2021) show that firing spikes at the 90-day tenure threshold, suggesting firms do in fact use the 90-day window to continue screening workers before committing to a permanent employment relationship. We take advantage of this labor market feature when testing model predictions for dismissal rates by referral status and race.

Brazil also has a large informal labor market. For us, the key distinction is between formal labor market contracts, which will appear in the administrative data, and informal contracts, which will

[11] In recent years, some state and municipal programs have adopted affirmative action policies, and some universities have begun to impose racial quotas in admissions (Francis and Tannuri-Pianto, 2013).

not. Over the period of our study, the informal sector accounts for between 40% and 60% of total employment, with the share declining over time. It is not uncommon for firms to employ some workers on formal contracts and others on informal contracts (Haanwinckel and Soares, 2020).

To provide more context, we summarize data from the Pesquina Nacional for Amostra de Domicilios (PNAD) between 2003 and 2015. Our discussion of summary statistics from the PNAD mirrors that of Gerard et al. (forthcoming). The PNAD is an annual, nationally representative household survey that collects information on labor market outcomes for both formal and informal workers. We limit to men and women ages 18–65. Statistics by race and gender are reported in Table 1.

[Table 1 about here.]

As shown in Table 1, about 48% of working-age Brazilian men and women are white, 43% are brown or mixed race, and 8% are black. This paper focuses on private sector employment. Thirty-nine percent and 21% of men and women work in the private sector, excluding the self-employed. These rates are similar across racial groups, though more variable among women. Unemployment rates are 25% to 30% higher for nonwhite men and women.

We next compare entrepreneurship rates by racial group. We define entrepreneurs as those who self-report running a formal or informal business with at least one paid employee. Overall, 3.0% of men and 1.6% of women are entrepreneurs. Entrepreneurship rates are more than twice as high among whites. For example, 4.1% of white men are entrepreneurs, while 2.1% and 1.8% of brown and black men are entrepreneurs.

Among private sector employees, white men and women have more years and education and receive wages that are 20 to 30 log points higher than wages received by nonwhite men and women.[12] About 80% of private sector employees report having a valid “carteira de trabalho,” which indicates that they are employed in the formal sector and hence are included in the RAIS data. Rates of formality are similar across racial groups.

3.2 RAIS Employer-Employee Data

Our analysis is focused on an extract of the RAIS data over the years 2003–2017. RAIS is a collection of administrative records reported by individual business establishments to the Brazilian labor ministry (*Ministerio do Trabalho* — MTE) for the primary purpose of administering various social security programs.

Each record captures the details of an employment contract between a worker and an establishment during a given year. The recorded details include the worker’s race, education, and gender as reported by the employer. The data also record contract-specific information including average monthly earnings over the year, occupation, the date of hire, and, for jobs that end, the date and cause of separation. We distinguish between employee-initiated separations (“quits”) and

[12]Gerard et al. (forthcoming) find that controlling for region, year, education, and experience reduces the wage gap between white and nonwhite private sector workers to 11%–13%.

employer-initiated separations ("dismissals"). The data include variables that identify both the individual establishment where an employee works and, separately, the firm or enterprise that owns the establishment.

[Table 2 about here.]

Table 2 shows significant racial disparities in wages. We limit the sample to worker-firm-year observations for men and women between the ages of 18 and 65 on private sector, indeterminate-length contracts for at least 30 hours per week. We report data separately for all worker-firm-year observations, recently hired workers (the first year of a job spell), and recently hired workers at entrant firms.[13] Within each category, we also compare characteristics of white and nonwhite workers. The full data are composed of 688 million job-year observations, of which 36.5% are for nonwhite workers. There is a 20 log point (22%) raw wage gap between white and nonwhite workers.[14] This may partially be explained by differences in characteristics: white workers are 7.3 percentage points more likely to be college graduates and 0.9 years older, on average. Recently hired workers are more likely to be nonwhite. The raw racial wage gap among new hires is, however, smaller than the overall gap, at 12.5 log points (13%). Racial differences between recently hired workers not significantly different when we limit the sample to entrant firms.

We must address a key issue in how race is recorded in RAIS. Cornwell et al. (2017) document that a non-trivial number of workers have different races reported by different employers in RAIS. This is possible because when a worker changes jobs, their new employer makes an independent record of their demographic characteristics. Cornwell et al. (2017) show that changes in reported race are not independent of residual changes in earnings and are not explained as simple misreporting. To address this issue, we identify the race for each individual using their modal reported race across all contract-years for which they appear in the data.

3.3 CNPJ Ownership Data

In one approach to inferring the race of a firm's founder, we follow Dias and Rocha (2021) and use publicly available data on firm ownership from the federal registry of firms, the *Cadastro Nacional de Pessoa Juridica* (CNPJ), maintained by the Receita Federal do Brazil.

The data report all individual and corporate owners with any stake in a firm. The publicly available data on firm ownership is limited to firms with more than one legal owner. For all individuals, the data include either the individual tax identifier (CPF) or a combination of name and a subset of the tax identifier. We use this identifying information to match individuals to the RAIS data. Hence, for all individual owners included in the CNPJ with some formal sector job spell from 2003–2017, we can identify the owner's race. We merge ownership data to firms in the RAIS data using the unique CNPJ firm identifier.

[13] We describe how we identify entrant firms in more detail in Section 4.2.

[14] We compute an hourly wage by deflating average monthly earnings by the product of contracted weekly hours and average weeks per month.

4 Testing Model Predictions

In this section we test the model predictions described in Section 2. In Section 4.1 we describe and implement tests that evaluate whether referral effects on hiring outcomes are declining in employer size and whether referral effects are larger for incumbent employees and connected job seekers of the same race. In Section 4.2 we test whether the racial composition of hires at firms with white and nonwhite founders converge as cumulative hires and employer size increase. In Section 4.3 we test predictions for differences in dismissal rates between referred and non-referred hires. In Section 4.4 we test predictions for dismissal rates by race of hire and employer characteristics.

4.1 Empirical Model of Referral Effects

If referrals are an important hiring channel, we should observe that firms are more likely to hire job seekers with a social connection to one of their incumbent employees. Conversely, job seekers should be more likely to move into firms where they have a social connection. Following Eliason et al. (2020), we evaluate the importance of referral hiring by modeling dyads that pair workers that change jobs from one year to the next with a set of potential destination establishments. In the model, i denotes a worker who is observed to separate from establishment j. The binary outcome variable $P_{ijk} = 1$ if i moves from origin establishment j to potential destination k. The variable of interest $C_{ijk} = 1$ if i has a social connection to some incumbent employee at k and is 0 otherwise.[15]

Our basic specification is a linear probability model for P_{ijk}:

$$P_{ijk} = \alpha_{jk} + X_{ij}\beta + \lambda C_{ijk} + \varepsilon_{ijk}. \tag{1}$$

where α_{jk} are fixed effects for employer origin-destination pairs. The parameter λ measures the increase in the probability that firm k hires a worker from origin j when that worker has a social connection to one of k's incumbent employees.

We measure social connections, C_{ijk}, using information on whether two workers have been coworkers in the past. Specifically, two workers are coworkers in our data if they were employed in the same establishment and the same occupation at the same time.[16]

4.1.1 Identification of Referral Effects

We use a novel combination of strategies to identify a referral effect in hiring from our information on past coworkers. In equation (1), the referral effect, λ, is identified under the assumption that coworker connections are random conditional on the employer origin-destination pairs. It therefore allows for arbitrary heterogeneity in the probability that k will hire a worker from j. The identifying

[15] We consider all workers who change jobs when constructing dyads for the analysis reported in this section. In Appendix B.2 we report results restricted to transitions involving workers from mass displacement events.

[16] We use eight top-level occupation codes from the 2002 vintage of Brazil's occupation classification system, the *Código Brasileiro de Ocupações* (CBO-2002). Following Eliason et al. (2020), we restrict attention to coworking relationships in plants with fewer than 100 employees. This restriction both helps to manage the size of the resulting data and to focus on environments where coworkers are likely to know one another.

assumption will be violated if it is not the connection per se, but instead something about the shared work history with an incumbent that makes k more likely to hire the linked worker. For example, if a restaurant manager hires a cook from another restaurant and learns they are well-trained in a particular type of cuisine, they may be more likely to hire another cook from that restaurant later.

To address this concern, we pursue a complementary identification strategy, comparing pairs of workers who worked together in the past with pairs that we call *placebo* coworkers: pairs that have held the same job, but not at the same time. In our data, we observe the months a worker is employed in any job. For any two workers that hold the same job in our data, we measure the number of months their employments spells overlap. When the pair were true coworkers, the overlap measure is positive. When they were not actual coworkers, the overlap is negative, and measures the number of months between their employment spells. We define placebo coworkers as those pairs of workers with overlap between -12 months and 0 months. We can then identify the connection effect on any outcome by comparing workers with true coworker links to the hiring firm and those with placebo links.

[Figure 1 about here.]

Figure 1 illustrates the contrast between true coworkers and placebo coworkers. The vertical axis measures the hiring share across dyads (multiplied by 100) restricted to those where either a true or placebo coworker link is present. The figure captures three findings relevant to this identification strategy. First, the discontinuous jump in the hiring share—around a 50% increase—between overlap values of 0 and 1 confirms that true coworking relationships have an effect on the likelihood of hiring beyond what can be explained by two workers having similar employment histories. Second, the hiring share has a strong positive gradient when overlap is positive but is flat when overlap is negative. This suggests that the relevance of actual coworking relationships to hiring outcomes is increasing in the amount of exposure two workers have to one another. Finally, the average value of the hire share is around 0.1 across all dyads (including those for which there is neither a true nor a placebo connection). The average is double that, around 0.2, for placebo connections, suggesting the placebo connections do in fact capture information relevant to the hiring outcome.[17]

Our augmented specification incorporating placebo coworker links is

$$P_{ijk} = \alpha_{jk} + X_{ij}\beta + \lambda C_{ijk} + \lambda^* A_{ijk} + \varepsilon_{ijk}, \tag{2}$$

where $A_{ijk} = 1$ if there is either a true or a placebo connection between worker i and an incumbent worker at k. Note that most dyads do not involve any connection, either true or placebo, so λ^* and λ can both be identified along with a constant term. However, we are primarily interested in

[17] Since our sample is based on all workers that change jobs, one might worry it is biased toward workers with especially productive social connections or toward workers employed in bad matches relative to their outside options. In Section B.2 we report the results for workers that separated during mass displacement events and find no evidence that these forms of bias are driving our results.

λ, which measures the increase in hiring associated with a true coworker connection relative to a placebo connection.

4.1.2 Data for the Referral Analysis

The results in this section support the key assumptions of our model. First, we document very strong referral effects that operate through past coworker connections. Second, our estimated referrals effects are almost entirely driven by coworker pairs that share the same race. Finally, referral effects are declining in the size of destination establishments.

We estimate referral effects using data covering all of Brazil between 2013 and 2017.

[Table 3 about here.]

Table 3 describes the sample, which is composed of 303,338,866 dyads.[18] Just 0.082 percent of dyads capture cases where worker i is hired by k. The transitioning worker has a true linked coworker connection in 4.1 percent of dyads. We define placebo worker links as described above. 8.4 percent of observations have either a true link or a placebo link. We report the size of the potential destination establishment in three groups. The majority of dyads (62.6%) involve potential destination firms in the smallest size group (1–99 workers). Of the remaining dyads, 21.7% involve potential destinations with 100–499 workers, and 18.9% with 500 or more workers. The underlying transitions cover 1,353,787 hired workers (column 2) connected to 9,216,640 incumbents (column 3). The shares of hired workers who are white and male are smaller than in the population overall, at 0.320 and 0.559, respectively. These differences reflect the non-randomness in who changes jobs. The demographic characteristics of connected incumbent workers are closer to the population of all workers. Note that the incumbents in column (3) measure incumbents from true coworker connections only and not from placebo links.

4.1.3 Baseline Referral Effects

[Table 4 about here.]

Columns 1, 2, and 3 of Table 4 report estimates of the effect of real coworker links, λ under different identifying assumptions. Column 1 reports estimates of equation (1), which includes establishment pair effects, but does not use the placebo link contrast, comparable to the main specification in Eliason et al. (2020). Under this model, we estimate λ =0.182, which is more than double the baseline value of 0.084 percent. In column 2, we eliminate establishment-pair effects (though we still control for separate origin and destination establishment effects) and include a control for "Any Link". In this specification, λ measures the true referral effect relative to

[18] Like Eliason et al. (2020), we restrict attention to jk pairs such that an incumbent worker in establishment k has a connection to a job mover in j. This restriction is without loss of generality since it is based on predetermined coworker relationships. In any case, λ is only identified from pairs of firms for which there is variation in the presence of a coworker link and in the outcome.

observations with a placebo link. With this specification, λ =0.222. The coefficient for "Any Link" measures the effect of placebo links and is identified relative to those dyads where there is neither a true coworker connection nor a placebo connection. Placebo links are associated with an increase in hiring of 0.097 percentage points; doubling the baseline.

Column 3 reports our preferred specification, based on (2), which uses both identification strategies. Under that model, the presence of a true coworker link increases hiring by 0.117 percentage points, which is 1.4 times the baseline mean. We also find a significant effect from placebo links, which increase hiring by 0.066 percentage points. Taken as a whole, our results support a substantial role for coworker links in hiring regardless of the specification. However, the data also show that firms are disposed to hire workers that have been employed in the same place as one of their incumbents even when they do not know each other.[19]

4.1.4 There is Racial Homophily in Referral Hiring

We now examine heterogeneity by race. Our model is predicated on the assumption of homophily—that referrals are much more likely between workers of the same race. While homophily is well-established in the literature, we document it in the Brazilian labor market. The empirical model investigates four possible pairings: nonwhite hired worker linked to a nonwhite incumbent; nonwhite hire linked to white incumbent, and so on. We interact indicators for these pairings with the indicator for true coworker links and with a variable indicating either a true or placebo link. The omitted category corresponds to dyads where there is neither a placebo nor a true coworker connection.[20]

Coworker effects are between seven and nine times stronger when coworkers are of the same race. For dyads in which a nonwhite job seeker is linked to a nonwhite incumbent, a true connection increases the hiring probability by 0.192, which is a 64 percent increase relative to the overall effect from column 3. When both workers are white, the coworker effect is 0.136. By contrast, the estimated effects are an order of magnitude smaller when the coworkers are of different races; just 0.025 when a nonwhite job seeker is connected to a white incumbent and 0.020 in the opposite case. These results support our premise that referrals are far more common between members of the same

[19]Section B.2 describes estimates of the same models on dyads for workers displaced from their origin employer. The results in the displaced workers sample are qualitatively identical, and very similar quantitatively, albeit marginally smaller; consistent with the arguments in Cingano and Rosolia (2012) and Caldwell and Harmon (2019) that some workers are drawn to change jobs because of the quality of their social networks. If this bias is indeed present, it has no economically relevant implications for the results in this section.

[20]More formally, we estimate an extension of equation (2):

$$\begin{aligned} P_{ijk} = \alpha_{jk} + X_{ij}\beta &+ \left[\lambda_{N,N} M_{ijk}^{N,N} + \lambda_{W,N} M_{ijk}^{W,N} + \lambda_{N,W} M_{ijk}^{N,W} + \lambda_{W,W} M_{ijk}^{W,W}\right] C_{ijk} \\ &+ \left[\lambda^*_{N,N} M_{ijk}^{N,N} + \lambda^*_{W,N} M_{ijk}^{W,N} + \lambda^*_{N,W} M_{ijk}^{N,W} + \lambda^*_{W,W} M_{ijk}^{W,W}\right] A_{ijk} + \varepsilon_{ijk}. \end{aligned}$$

The indicator $M_{ijk}^{W,N}$ takes a value of one when the hired worker, i, is white and the incumbent worker to whom they are connected in firm k is nonwhite. The coefficient $\lambda^{W,N}$ measures the strength of the referral effect for this type of pairing relative to the value for placebo connections. The other indicators and coefficients are defined and interpreted similarly. As in (2), A_{ijk} indicates the presence of any link, either real or placebo.

race. Moreover, because white and nonwhite workers tend to work in different places, same-race links are much more common, suggesting that our results understate the role of homophily in referral hiring.

[Figure 2 about here.]

Figure 2 illustrates the relationship between race match and hiring using more detailed information on overlap. As in Figure 1, overlap measures the number of months a job seeker was employed in the same prior job as the linked incumbent. Negative values indicate the number of months between job spells. The discontinuity between non-positive and positive levels of overlap shows the importance of actual social interactions in hiring. The largest discontinuities are for Nonwhite / Nonwhite and White / White coworker pairs. The data support a mild role for social interactions where white job seekers are linked to non-white incumbents, but these are much smaller. While it is not directly relevant for our analysis of referrals, for negative values of overlap there is still a large hiring effect for nonwhite coworker pairs. One explanation is that firms associate the quality of their nonwhite workers with their prior employers and are more likely to hire nonwhite workers when they come from firms that have provided other successful workers in the past.

4.1.5 Referral Hiring Is Declining in Employer Size

[Figure 3 about here.]

Finally, we assess the relationship between referral use and employer size. Figure 3 is consistent with our assumption that large firms are less likely to use referrals. The figure plots referral effects estimated from a version of (2) that allows for heterogeneity according to the size of the hiring job.[21] The plotted effects measure the proportional increase in the probability a worker is hired when a linked coworker is present. With one exception, the effect of a true coworker link effects is monotonically decreasing. In establishments with fewer than four workers, a true coworker link increases the hiring probability by a factor of 2.52 relative to the mean for that size group. For firms with over 1,000 workers, the estimated effect is an increase of just 0.7 times the mean.

These findings are consistent with larger establishments having more formalized human resources (HR) practices. This relationship is corroborated in the Brazilian wave of the World Management Survey (WMS), which scores firms on their adoption of formal management practices across several domains, including people management, operations, and performance targeting (Bloom et al., 2014). Appendix Figure B.1 shows that the adoption of formal people management

[21] We estimate:

$$P_{ijk} = \alpha_{jk} + X_{ij}\beta + \left[\lambda + \sum_s \delta_s \mathbb{1}(S_k = s)\right] C_{ijk}. + \left[\lambda^* + \sum_s \delta^*_s \mathbb{1}(S_k = s)\right] A_{ijk} + \varepsilon_{ijk}.$$

where, S_k indicates the size class of destination plant k. We report the effect magnitude as $\left(\hat{\lambda} + \hat{\delta}_s\right) / \overline{P}_s$, with $\overline{P}_s$ measuring the average outcome for dyads with destinations in size class s. The inclusion of A_{ijk} means the effects are identified relative to the value for placebo connections.

and overall management practices are increasing in firm size.[22] The WMS for Brazil only covers a small sample (815 observations) on medium-sized manufacturing firms. However, we can provide complementary evidence that the adoption of formal HR management patterns is increasing in employer size more generally. For all firms and establishments in RAIS, we proxy for HR formality using the share of an establishment's employees in HR-related occupations.[23] Appendix Figure B.2 plots the HR share of an employer's workforce by employer size, where employers are defined at either the establishment or firm level. The relationship is increasing for either measure.

4.2 Racial Composition of Hires Converges with Cumulative Hires and Size

We predict that the racial composition of new hires for an employer will be correlated with the race of the founder, but this correlation is decreasing in the employer's (a) cumulative number of hires and (b) size (n from the model). As the employer's cumulative number of hires and size increase, new hires are further removed from the founder's referral network. Together, these predictions imply that, in the cross-section, the nonwhite share of hires is increasing in cumulative hires for firms with white founders and is decreasing in cumulative hires for firms with nonwhite founders. We first test for this pattern in the cross-section and then test whether the pattern holds within firms and between firms with more or fewer total hires, which roughly corresponds to predictions (a) and (b).

We take all firms that we observe as entrants in the RAIS data. For multi-establishment firms, we take the first establishment observed for the firm, if we observe that establishment's year of entry. We refer to these establishments and single-plant firms as *headquarter (HQ)* establishments (we refer to HQ establishments and firms interchangeably for the remainder of the paper). We further restrict to establishments that enter the RAIS data with 1–49 employees as of December 31 in its year of entry. We are left with a sample of about 2.3 million HQ establishments. Appendix Table B.3 provides descriptive statistics for these entrant HQ establishments.[24]

We characterize founder race in two ways. First, following standard practice in the entrepreneurship literature (Kerr and Kerr, 2017; Azoulay et al., 2020; Babina, 2020; Bernstein et al., 2021), we infer the race of a firm's founder using the race of the highest paid manager in the HQ establishment at entry.[25] Second, when possible, we infer the race of a firm's founder using the racial composition of ownership. We classify firms as having a white founder when we can identify more than 50%

[22] Cornwell et al. (forthcoming) show that the positive relationship between employer size and the WMS people management score continues to hold when conditioning on other observable characteristics.

[23] This includes the following occupations: *administrador* (administrator), *diretor de recursos humanos* (human resources director), *gerente de recursos humanos* (human resources manager), and *gerente de departamento pessoal* (personal department manager).

[24] Our definition of entrant firms includes preexisting informal firms that formalize, a category that we are unable to separately identify.

[25] For HQ establishments with no employee with a manager occupation code, we take the highest paid employee. If multiple people have the same exact wage at the top of the distribution, we pick one randomly. Using tax data on S corporations in the United States, Azoulay et al. (2020) find 90% of owner-workers are among the top three earners in the firm during the first year. Note that while this procedure may not identify the relevant founder in some cases, the race of the individual identified by this procedure is likely highly correlated with the race of the founder.

of ownership as white and as having a nonwhite founder when we can identify more than 50% of ownership as nonwhite.[26] Using either classification, we find that entrant firms with white and nonwhite founders are similar in terms of their size and survival rates (see Appendix Table B.3).

We then ask how the composition of new hires in subsequent years evolves with the establishment's cumulative number of hires. For firms with a white founder, we predict the nonwhite share of hires to be increasing in cumulative hires. For firms with a nonwhite founder, we predict the nonwhite share of hires to be decreasing in cumulative hires. We predict that the nonwhite share of hires for HQ establishments of firms with white and nonwhite founders will converge as their cumulative hires increase so that when cumulative hires is sufficiently high, the racial composition of hires is not related to the race of the founder.

We estimate regression models of the form

$$\begin{aligned}\log(E(\text{NONWHITE}_{it}|\cdot)) = &\sum_n \sum_r \eta^{n,r} \times \mathbb{1}_{\{N(J,t)=n\}} \times \mathbb{1}_{\{R(J)=r\}} \\ &+ \tau_t + \mu_{m(J(i,t))} + \omega_{o(i,t)} + \epsilon_{it} \end{aligned} \tag{3}$$

via Poisson quasi maximum likelihood (Correia et al., 2020), where each observation is a new hire, i indexes workers, t indexes time, and $J(i,t)$ indexes the establishment.[27] We limit to hires made after the year of the firm's entry. NONWHITE_{it} is an indicator for whether the new hire is nonwhite. τ_t are year fixed effects, $\mu_{m(J(i,t))}$ are microregion fixed effects (which we use to approximate local labor markets), and $\omega_{o(i,t)}$ are fixed effects for two-digit occupation. $R(J)$ categorizes establishments by founder race. $N(J,t)$ indexes an employer's cumulative hires to date. We group hires into increments of five: hires 1–5, 6–10, 11–15, and so on. The omitted category is the first increment of hires for establishments with white founders. The $\eta^{n,r}$ coefficients have a natural proportional interpretation: they measure the proportional increase in the probability that a hire is nonwhite relative to the omitted category.

We plot the coefficient estimates in Panel A of Figure 4. Here we infer founder race from the race of the top-paid manager. (We plot analogous results where we infer founder race using the racial composition of ownership in Appendix Figure B.4; the results are similar.) The pattern fits our predictions. For early hires, the racial composition of new hires is closely tied to founder race. For the first few hires, in Panel A the probability that the hire is nonwhite is about 35 log points higher at firms with a nonwhite founder compared to firms with a white founder. This gap declines steeply in cumulative hires. By the 50$^{\text{th}}$ hire, this gap declines to about 15 log points, and 5 log points by the 200$^{\text{th}}$ hire. By the 350$^{\text{th}}$ hire, the racial composition of hires at firms with white and nonwhite founders is statistically indistinguishable. Interestingly, the convergence in racial

[26] Given the racial disparities reported in household survey data as described in Section 3, the first method may inflate the nonwhite share of founders but covers a substantially larger set of firms. Note that in calculating the white and nonwhite share of ownership, we include owners that we cannot match to the RAIS data in the denominator.

[27] Note that the fixed effects Poisson estimator only invokes the conditional mean assumption in (1) and a standard strict exogeneity assumption. It is well suited to binary outcomes and does not require that the data follow a Poisson distribution. See Wooldridge (1999). We have also estimated λ under the assumptions of a linear probability model, as in Eliason et al. (2020), and obtain similar results.

composition is driven entirely by firms with white founders. Among firms with nonwhite founders, the relationship between cumulative hires and the racial composition of new hires is essentially flat.

[Figure 4 about here.]

The pattern illustrated in Panel A of Figure 4 may both reflect the within-firm evolution in the racial composition of hires and differences in composition for firms with more or fewer total hires. The model predicts that both channels will contribute to the cross-sectional relationship. We next distinguish between these channels.

We test for whether the nonwhite share of hires varies as predicted both *within* establishments and *between* establishments. We estimate equation (3) but allow the η coefficients to vary with a firm's total observed hires. Specifically, we estimate

$$\log(E(\text{NONWHITE}_{it}|\cdot)) = \sum_s \sum_n \sum_r \eta^{s,n,r} \times \mathbb{1}_{\{S(J)=s\}} \times \mathbb{1}_{\{N(J,t)=n\}} \times \mathbb{1}_{\{R(J)=r\}} + \tau_t + \mu_{m(J(i,t))} + \omega_{o(i,t)} + \epsilon_{it}, \quad (4)$$

where $S(J)$ categorizes firms by their total observed hires: 50–249, 250–499, and 500+. $N(J,t)$ now groups hires into increments of 10. We restrict estimation to hires 1–50 for firms with 50–249 total observed hires, hires 1–250 for firms with 250–499 total observed hires, and hires 1–500 for firms with 500+ total observed hires. This restriction maintains a balanced sample of firms contributing to the estimation of $\eta^{s,n,r}$ coefficients.

The results are presented in Panel B of Figure 4. For establishments with white founders, the pattern of coefficients is similar to what we found in Figure 4, though the magnitude of change is smaller. The relationship between an establishment's nonwhite share of hires and cumulative hires is increasing and concave. By the 200$^{\text{th}}$ hire, the probability that the hire is nonwhite is about 20 log points larger than that same probability for the first hire. For establishments with nonwhite founders, the relationship is negative rather than flat, but the magnitude of change is smaller than that for establishments with white founders.

Panel B of Figure 4 also illustrates clear between-firm differences by total hires. Among firms with white founders, the nonwhite share of hires is increasing in total observed hires for all values of cumulative hires. For initial hires, the likelihood that a hire is nonwhite is about 8% greater at firms with 250 or more observed hires than at firms with 50–249 observed hires. This pattern is consistent with the premise that larger firms hire a smaller share of their workforce via referral.

The pattern is somewhat different for firms with nonwhite founders. For these firms, the racial composition of initial hires is unrelated to firms' total observed hires. For later hires, the nonwhite share of hires is *decreasing* in total observed hires. As a result, the speed of convergence (as measured by cumulative hires rather than time) is faster for firms with fewer observed hires. This may reflect that new hires are slower to generate referrals at firms with more observed hires because, for example, there is less time between hires at these firms.

We extend the analysis in three ways in the Appendix. First, we conduct an analogous exercise for new establishments that are subsidiaries of existing firms. We characterize establishments by the racial composition of the firm's incumbent employees. The findings are similar (see Figure B.6). Early on, establishments from firms with mostly white employees are more likely to hire white workers than peer establishments from firms with mostly nonwhite employees. But these differences disappear as the establishment's cumulative hires increase.

The next two extensions examine whether the patterns documented here vary with other firm characteristics. Motivated by Gerard et al. (forthcoming), who find that nonwhite workers are underrepresented at high-paying firms, we estimate equation (3) separately by firm pay premium quintile, where we estimate pay premiums using the canonical two-way fixed effects model of Abowd et al. (1999). We find similar patterns for low- and high-paying firms (see Appendix Figure B.7).

Finally, we test for heterogeneity by the racial composition of the local labor market. We divide firms into quintiles by the nonwhite share of hires in their microregion and estimate equation (3) separately by quintile. We find similar patterns for firms in local labor markets with large white and nonwhite majorities (see Appendix Figure B.8).

4.3 Referral Hiring and Learning About Match Quality

A common explanation for why employers use referral networks in hiring is that they can obtain more information about the match quality of potential referral hires (Simon and Warner, 1992; Topa, 2019). A growing literature tests the empirical implications of this class of referral-based job search models (Dustmann et al., 2016; Brown et al., 2016). These papers test whether, within a firm, referral hires have lower turnover relative to non-referral hires, and whether these differences dissipate with tenure.

We first check whether connected hires have lower dismissal rates than non-connected hires and how differences in dismissal rates evolve with tenure. We estimate a discrete-time hazard model and compare hazard rates within the same establishment and occupation for hires who are connected and not connected to an incumbent employee at their time of hiring. We expand our job spells data into a job spell by time period data set, where each observation represents a job spell and 15-day tenure period, where periods are indexed by p.[28]

We estimate regression models of the form

$$\log(E(\text{DISMISSED}_{iJ(i,t)p}|\cdot)) = \sum_p \theta^p \text{CONNECTED}_{iJ(i,t)} \mathbb{1}_p + \sigma_p + \tau_t + \omega_{o(i,t)} + \psi_{J(i,t)} + \epsilon_{it}, \quad (5)$$

where $\text{DISMISSED}_{iJ(i,t)p}$ is indicator for whether the establishment $J(i,t)$ dismisses employee i in tenure period p, $\text{CONNECTED}_{iJ(i,t)t}$ is an indicator for whether hire i has a connection at establishment $J(i,t)$ at the time they are hired, and $\psi_{J(i,t)}$ are establishment fixed effects. We limit estimation to establishment-by-occupation cells with at least one connected hire and one

[28] We are missing data on the specific day of separation for the years 2011–2013. For this reason, we exclude job spells that begin in 2009–2013 from the analysis.

non-connected hire. The coefficients θ^p convey the differences in log dismissal hazard rates in tenure period p between connected and non-connected hires conditional on year, occupation, and establishment fixed effects.

Panel A of Figure 5 plots dismissal rates by tenure for connected and non-connected hires as implied by equation (5). There are two points to note. First, dismissal rates for non-connected hires exceed dismissal rates for connected hires at all depicted tenure levels. This is particularly salient in the period running up to the end of the 90-day probationary period, where dismissal rates spike. While both non-connected and connected hires experience a sharp increase in dismissal rates, the spike is markedly larger for non-connected hires. Second, the gap in dismissal rates generally dissipates over time following the 90-day spike, with much of the closing in the gap occurring after about 250 days on the job.

[Figure 5 about here.]

To provide additional evidence that referral hires are less likely to be dismissed, we compare outcomes for connected hires and placebo connected hires as in Section 4.1. For each hire i at establishment $J(i,t)$, we denote the maximum overlap in prior job spells across combinations of worker i and incumbent workers at establishment $J(i,t)$ by $\text{OVERLAP}_{iJ(i,t)}$. As discussed in Section 4.1, a negative overlap between a pair of workers indicates that the two were not true coworkers, and the value indicates the number of months between their job spells. In cases where hire i does not share a prior job with any incumbent in establishment $J(i,t)$, we set $\text{OVERLAP}_{iJ(i,t)} = -\infty$.

We focus on dismissals during the probationary period. We estimate regression models of the form

$$\log(E(\text{DISMISSED-3M}_{it}|\cdot)) = \sum_{k\in\mathcal{K}} \theta^k \mathbb{1}_{\text{OVERLAP}_{iJ(i,t)}\in k} + \tau_t + \omega_{o(i,t)} + \psi_{J(i,t)} + \epsilon_{it}, \tag{6}$$

where DISMISSED-3M_{it} is an indicator for dismissal within three months of the hire date and $\mathcal{K}$ categorizes $\text{OVERLAP}_{iJ(i,t)}$ values as follows: -12 to -9, -8 to -6, -5 to -3, -2 to 0, 1 to 3, 4 to 6, 7 to 9, 10 to 12, 13 to 15, 16 to 18, and > 18. The omitted category is < -12 (including $-\infty$). Hence, the θ^k coefficients identify the dismissal rates of placebo connected and connected hires as a function of overlap relative to hires who are neither.

Placebo connected hires are less likely to be dismissed than hires who are neither connected nor placebo connected. Among placebo connected hires, the θ^k coefficients are consistently around –1.25, indicating that hires who are neither connected nor placebo connected are about three times more likely to be dismissed during the probationary period than their placebo connected peers. Among placebo connected hires, the relationship between overlap and dismissal rates is flat. There is a clear trend where overlap is greater than zero; among connected hires, dismissal rates are *decreasing* in the length of time their prior job spell overlapped with an incumbent employee.

Note that we only capture one type of social connection in our data, previous coworkers, and turnover patterns may vary with the type of connection. Nonetheless, we interpret this striking pattern as evidence that the referral hires have lower dismissal rates than comparable non-referral hires.

Overall, our findings are consistent with job search models where referral networks provide information to employers about a job candidate's match quality. If indeed the hiring dynamics documented in Section 4.2 are driven by referral hiring, racial differences in dismissal rates for recent hires favor hires of the same race as the establishment's founder but are decreasing in an establishment's total number of hires.

4.4 Racial Disparities in Dismissal Rates Are Decreasing in Total Hires

We test whether within-firm racial differences in dismissal rates are decreasing in total hires. We return to our sample of entrant firms and estimate regression models of the form

$$\log(E(\text{DISMISSED-3M}_{it}|\cdot)) = \tau_t + \omega_{o(i,t)} + \psi_{J(i,t)} + \psi^{NW}_{J(i,t)} + \epsilon_{it}, \tag{7}$$

where $\psi_{J(i,t)}$ are firm fixed effects and $\psi^{NW}_{J(i,t)}$ are firm by nonwhite fixed effects. Hence, $\psi^{NW}_{J(i,t)}$ is the firm-specific racial disparity in log three-month dismissal rates.

Figure 6 depicts the average of $\psi^{NW}_{J(i,t)}$ as a function of a firm's total observed hires and race of founder. We limit to firms where we observe at least 20 hires and then take an average of $\psi^{NW}_{J(i,t)}$ across firms, weighting by each firm's number of observed hires. At firms with 20–49 hires and white founders, the three-month dismissal rate is about 18% higher for nonwhite hires. This declines to a 5% gap at firms with 500 or more hires. By contrast, at firms with 20–49 hires and nonwhite founders, the three-month dismissal rate is about 5% *lower* for nonwhite hires. There is essentially no racial difference in dismissal rates at firms with nonwhite founders and 250 or more hires.[29]

[Figure 6 about here.]

If the pattern shown in Figure 6 is driven by the fact that racial differences in the referral share of hires is diminishing with total hires, then the relationship between racial disparities in dismissal rates and total hires should be muted with tenure. We test this by re-estimating equation (7) but replacing the outcome with an indicator for dismissal in the first 18 months of the spell. We plot the corresponding coefficients also in Figure 6. Reassuringly, we find that the relationship between total hires and the racial disparity in 18-month dismissal rates is relatively flat.

5 Alternative Interpretations

We interpret the evidence presented in the previous section as consistent with the job search model presented in Section 2. In this section we consider alternative interpretations of the patterns we document.

[29] Appendix Figure B.10 shows similar patterns for quits and all separations.

5.1 Human Capital

One alternative explanation for our findings is that they are driven by differences between white and nonwhite workers in human capital or preferences over occupations. Firms with white and nonwhite founders may hire for positions that tend to be filled by nonwhite and white workers later in the firm's life cycle, respectively. Equation (3) includes fixed effects for two-digit occupation; however, the positions that firms fill may vary in unobservable ways over the firm life cycle.

To assess this explanation, we examine how the occupational composition of hires varies over the firm life cycle for firms with white and nonwhite founders. For each *six*-digit occupation, we measure the nonwhite share of workers hired into that occupation, $\bar{\omega}_o$. We then estimate models analogous to equation (3), replacing the outcome with $\bar{\omega}_o$ and replacing microregion fixed effects with establishment fixed effects:

$$\log(E(\bar{\omega}_{o(i,t)}|\cdot)) = \sum_n \sum_r \eta^{n,r} \times \mathbb{1}_{\{N(J,t)=n\}} \times \mathbb{1}_{\{R(J)=r\}} + \tau_t + \psi_{J(i,t)} + \epsilon_{it}. \tag{8}$$

Appendix Figure B.11 plots the $\eta^{n,r}$ coefficient estimates. The value of $\bar{\omega}_{o(i,t)}$ is increasing slightly over the firm's life cycle. However, this is true for both firms with white and nonwhite founders. Racial differences in human capital do not appear to explain our findings.

5.2 Taste-Based Discrimination

A second alternative explanation is that employers exhibit taste-based discrimination. Some establishments with white founders may prefer to employ white workers conditional on match productivity and are less likely to grow as a result. This could generate an increasing relationship between cumulative hires and the probability that a hire is nonwhite, as documented in Section 4.2.

There are two patterns we show that are inconsistent with at least a standard employer-driven taste-based discrimination model. First, the convergence in the racial composition of hires with cumulative hires holds *within establishment*. A model where employer tastes over employee race are fixed would not generate this result. Second, a standard taste-based discrimination model would not generate the finding that there are racial differences in dismissal rates that dissipate with cumulative hires. If racially biased employers set a lower threshold for dismissing hires from their disfavored group, then a forward-looking employer would account for this preference at the hiring stage and set a more demanding threshold for hiring job seekers from that group. It may be possible to rationalize this pattern with a more complicated taste-based model where decision-makers who make hiring and firing decisions are different and misaligned (Lehmann, 2013).

There is also a sense in which taste-based discrimination may play an implicit role in our model. Incumbent employees have discretion over which job seekers to refer, and over which social connections to form in the first place.

5.3 Worker Preferences

A thrid interpretation is that the patterns we show reflect job seeker preferences over workplace characteristics. In particular, nonwhite job seekers prefer to not work at small or young employers with white founders. To evaluate this alternative hypothesis, we build on the insight that, under some assumptions, worker preferences over employers can be inferred from worker mobility patterns (e.g., Sorkin, 2018).

There are now several approaches to constructing a revealed preference ranking of employers. We use the *poaching rank* as developed by Bagger and Lentz (2018). The premise of the poaching rank is that higher-ranked employers should hire relatively more workers from employment than from unemployment. That is because poaching a worker from another employer indicates the worker prefers the destination employer. The poaching index for establishment J is defined as the share of all new hires who are poached from other employers:

$$p_J = \frac{n(.,J)}{n(0,J) + n(.,J)}, \tag{9}$$

where $n(.,J)$ is the number of hires poached from other and $n(0,J)$ hires from unemployment. The poaching rank of establishment J is a conversion of the poaching index into an ordinal ranking of employers. We present more details on the construction of the poaching rank in Section B.1.

In Appendix B we examine how race-specific poaching indices vary with founder race and total hires (see Appendix Figure B.12). We do not find evidence that nonwhite job seekers prefer to not work at small or young employers with white founders.

5.4 Complementarities in Production

A fourth alternative explanation is that workers are more productive when their coworkers are of the same race (Lang, 1986). Complementarities in production would naturally lead to workplace segregation. The racial composition of hires could converge across firms if these complementarities are stronger among early employees. Without productivity data, this explanation is difficult to rule out. However, it is not clear why this mechanism would generate the racial differences in dismissal rates that we observe.

5.5 Screening Ability

A fifth interpretation is that firms with white and nonwhite founders are better at screening white and nonwhite job seekers, respectively, even when those job seekers are not referrals. This would be consistent with suggestive evidence that managers are better at screening applicants who share their racial or ethnic background (Giuliano et al., 2011; Åslund et al., 2014; Benson et al., 2019).[30] This alternative story can match the findings that firms with white founders are more likely to

[30] Note, however, that this research generally cannot rule out that managers' superior screening of applicants who share their background is driven by referral hiring.

hire white employees than comparable firms with nonwhite founders and that firms with white and nonwhite founders are more likely to dismiss their nonwhite and white hires, respectively.

This mechanism could also generate convergence between firms with white and nonwhite founders if, for example, investments in formal HR practices or employer learning reduces the same-race advantage in screening ability.

We view this mechanism as similar to referral hiring in that they both posit that social (or cultural) proximity affects screening ability. We focus on referral hiring because aspects of referral hiring that are essential to the model examined here—-that referral networks are racially segregated and employers have additional information about the match quality of referral candidates—-are empirically supported in our setting and the broader literature. By contrast, the idea that managers are better at screening same-race applicants and that this advantage diminishes as cumulative hires increase is more speculative.

6 Implications for Racial Inequality

We have documented evidence that, on average, employers screen job seekers with more precision when they share the founder's racial background. This advantage in screening precision is declining in an employer's cumulative hires and size. We also discuss in Section 3 that entrepreneurship rates for white adults are twice as high as the same rates for nonwhite adults in Brazil.[31] In combination with our findings, this suggests that referral hiring will disadvantage nonwhite job seekers in the aggregate by making it more difficult for them vacancies where they are well matched (for example, see Bolte et al., 2020).[32]

In this section we argue that the dynamic effects of referral hiring, combined with racial differences in entrepreneurship, can help explain three stylized facts about racial differences in labor market outcomes. First, nonwhite workers are more likely to be dismissed by their employers than white workers. Second, nonwhite workers have less seniority than their white coworkers, on average. Third, nonwhite workers sort to larger employers than white workers.

6.1 Dismissal Rates

In our sample, 17.6% and 15.8% of nonwhite and white hires are dismissed within 90 days of their start date. In the United States, black workers are more likely than white workers to be laid off, fired, or discharged (Lang and Lehmann, 2012; Cavounidis et al., 2021). In the context of our model, these racial differences in "involuntary" separations could be explained by the fact that nonwhite or black hires are less likely to be referral hires.

[31] Racial disparities in entrepreneurship are not limited to Brazil. In the United States, about 13% of the adult population was black in 2012, while only 2% of businesses with at least one paid employee were black owned (Camara et al., 2019).

[32] Just as groups facing employer discrimination in Becker (1957) can avoid its ill effects on wages if the marginal employer is not discriminatory, the degree that referral hiring affects racial inequality in labor market outcomes may depend on the ability of nonwhite job seekers to sort to employers where they are socially connected and able to find well-matched positions.

To assess whether referral hiring can explain aggregate differences in dismissal rates, we compare dismissal rates in firms with white and nonwhite founders. We document in Section 4.4 that within firms with white founders, white hires are less likely to be dismissed than nonwhite hires. The opposite is true for firms with nonwhite founders, though the magnitude of racial differences in dismissal rates are smaller. Here we conduct a similar but distinct exercise. While in Section 4.4 we measure *within-firm* gaps in dismissal rates, coinciding with our employer-level model and model predictions, here we pool firms and combine both within-firm and between-firm variation in dismissals.

We estimate the following model, separately for all entrant firms, firms with white founders and firms with nonwhite founders:

$$\log(E(\text{DISMISSED-3M}_{it}|\cdot)) = \tau_t + \omega_{o(i,t)} + \beta\text{NONWHITE}_i + \epsilon_{it}. \tag{10}$$

Estimates for equation (10) are presented in Panel A of Table 5. Columns 1–3 pool all entrant firms, columns 4–6 limits to firms with white founders, and columns 7–9 limit to firms with nonwhite founders. Columns 1, 4, and 7 include only year fixed effects as additional controls; columns 2, 5, and 8 include year fixed effects and occupation fixed effects; and columns 3, 6, and 9 include year fixed effects and education fixed effects. For all establishments pooled, nonwhite hires are dismissed at an elevated rate. Without adjusting for job or worker characteristics (column 1), nonwhite hires are about 8% more likely to be dismissed within three months. This declines to 4% or 8% with the inclusion of occupation or education fixed effects.

[Table 5 about here.]

These differences are driven primarily by firms with white founders, where nonwhite hires are 8%–12% more likely to be dismissed, depending on the specification. By contrast, at establishments with nonwhite founders, nonwhite hires are 1% to 4% more likely to be dismissed.

6.2 Seniority

Buhai et al. (2014) find that separation rates are decreasing and wages are increasing in *seniority*, defined as worker's tenure relative to the tenure of their colleagues. Our model as written does not have substantive predictions for seniority because we ignore quits. However, the logic of the model suggests that nonwhite employees at establishments with white founders will tend to have less seniority than their white coworkers because they are hired later in an establishment's life cycle.

Following Buhai et al. (2014), we define a worker's *seniority index* as follows. Define q_{ijt} as the number of workers in establishment j with tenure greater than or equal to tenure of worker i at time t. Define n_{jt} as the total number of workers in establishment j at time t. The seniority index is defined as

$$\log r_{ijt} \equiv \log n_{jt} - \log q_{ijt}. \tag{11}$$

We estimate the following linear model, separately for all entrant firms, firms with white founders and firms with nonwhite founders:

$$\log r_{ijt} = \tau_t + \omega_{o(i,t)} + \beta \text{NONWHITE}_i + \nu \log n_{jt} + \epsilon_{it}. \tag{12}$$

Estimates for equation (12) are presented in Panel B of Table 5. Columns 1–3 pool all entrant firms, columns 4–6 limit to firms with white founders, and columns 7–9 limit to firms with nonwhite founders.

Overall, nonwhite employees have 4% to 5% less seniority than white employees. This is driven by firms with white founders, where nonwhite employees have 9% to 10% less seniority. At firms with nonwhite founders, nonwhite and white employees have similar seniority, on average.

6.3 Employer Size

Referral hiring can explain a striking pattern present in both Brazil and the United States that has received little attention: nonwhite and black workers sort to larger employers (Holzer, 1998; Miller, 2017). The sorting of nonwhite and black workers to large employers is perhaps surprising given that (1) large employers tend to pay more and employ more educated workers (Brown and Medoff, 1989) and (2) nonwhite workers tend to work at lower-paying firms, at least in Brazil (Gerard et al., forthcoming). We show that other observable job characteristics cannot explain this sorting pattern.

To characterize the relationship between the racial composition of new hires and establishment size, we estimate models of the form

$$\log(E(\text{NONWHITE}_{it}|\cdot)) = \sigma_{c(J(i,t))} + \tau_t + \mu_{m(J(i,t))} + \omega_{o(i,t)} + \epsilon_{it}, \tag{13}$$

where $\sigma_{c(J(i,t))}$ are fixed effects for the establishment's size category and $\omega_{o(i,t)}$ are fixed effects for combinations of occupation, industry, and worker education. Establishment size is measured as of December 31 in the year of the hire.

The estimated $\sigma_{c(J(i,t))}$ coefficients for various specifications of equation (13) are presented in Figure 7. The omitted category is establishments with one to four employees. The first specification includes only year fixed effects as additional explanatory variables. The coefficient of 0.044 for establishments with five to nine employees indicates that the nonwhite share of hires is 4.4 log points (4.5%) larger at establishments with five to nine employees relative to establishments with one to four employees. Coefficients are monotonically increasing in establishment size. The coefficient is 0.351 for establishments with 1,000 or more employees, indicating that the nonwhite share of hires is 42% larger at these establishments relative to establishments with one to four employees.

[Figure 7 about here.]

We increase the saturation of the model with each specification. The second specification includes microregion fixed effects, to little effect. Large establishments are not disproportionately concentrated in local labor markets with large nonwhite populations. The third specification includes three-digit occupation by two-digit industry fixed effects. These fixed effects are rich characterizations of jobs; they explain 52% of variation in starting log wages. Including occupation by industry fixed effects reduces the magnitude of the size effects moderately. For example, the coefficient for establishments with 1,000 or more employees declines from 0.351 to 0.249. However, a robust size gradient remains. The fourth specification replaces occupation by industry fixed effects with occupation by industry *by worker education* fixed effects, where worker education is divided into three categories: less than high school education, high school graduate, and college graduate. Incorporating education slightly *increases* the magnitude of the $\sigma_{c(J(i,t))}$ coefficients.

We conclude that nonwhite workers sort to larger establishments and this pattern cannot be explained by other job characteristics potentially correlated with size, including location, occupation, industry, and educational requirements.

7 Conclusion

We present a simple job search model with referral hiring, test its predictions using administrative data from Brazil, and consider its consequences for racial inequality. We emphasize the implications of referral hiring for how the racial composition of an employer's hires varies (1) with the race of the founder, (2) over time, and (3) with the employer's size. Among other predictions of the model, we confirm that (a) firms with white and nonwhite founders are more likely to hire white and nonwhite employees, (b) these differences disappear as establishments' cumulative number of hires increases, (c) firms are less likely to dismiss recent hires of the same race as the founder, and (d) racial differences in dismissal rates are also decreasing in an establishment's cumulative number of hires.

Given substantial racial disparities in entrepreneurship rates, the widespread practice of referral hiring appears to disadvantage nonwhite workers (relative to white workers) in the aggregate and to be a source of what sociologists refer to as "institutional discrimination" (Small and Pager, 2020). We show that referral hiring can help explain why nonwhite workers are more likely to be dismissed from their jobs, have less seniority, and sort to larger employers. In particular, racial differences in dismissal rates and seniority are driven by firms with white founders. Though we cannot identify workers' social connections directly, our findings suggest that, compared to their white peers, nonwhite job seekers may be connected to fewer firms and connected to larger firms, which tend to be less dependent on referral hiring. Though beyond the scope of this paper, a natural open question is what implications the combination of referral hiring and racial differences in entrepreneurship have for racial inequality in wages and employment rates.

We note two implications that our findings have for policy. First, our findings provide a novel rationale for affirmative action policies. Over the course of a firm's life cycle, the racial composition of its hires converges to the composition of the external market. But this convergence is slow in that few firms reach the scale where founder race no longer predicts a firm's racial composition of hires. Our findings suggest that a temporary affirmative action policy would accelerate this process by incentivizing firms to hire workers from groups underrepresented at the firm relative to the external market. Such an intervention would have short run costs—for example, an increase in dismissal rates or investments to improve the screening of external market candidates (Miller, 2017)—but would lead to persistent reductions to racial inequality in labor demand.

Second, our findings suggest that market frictions that affect the size distribution of firms will have implications for racial inequality in the labor market (Restuccia and Rogerson, 2017). For example, if small, productive firms are unable to expand to their efficient size due to some resource misallocation, these firms are also less likely to reach the point of having a racially diverse workforce. The logic of our findings suggests that the aggregate costs of misallocation will be disproportionately borne by groups underrepresented among entrepreneurs. On the other hand, dynamic markets with high firm turnover may also disadvantage groups underrepresented among entrepreneurs if firms that reach the scale needed to employ a diverse workforce make up a small share of the market.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–334.

_, Robert H. Creecy, and Francis Kramarz, "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data. *Cornell University Mimeo.*," 2002.

Aigner, Dennis J. and Glen G. Cain, "Statistical Theories of Discrimination in Labor Markets," *ILR Review*, January 1977, *30* (2), 175–187.

Arcard, Jean-Louis and Beatrice d'Hombres, "Racial Discrimination in the Brazilian Labour Market: Wage, Employment and Segregation Effects," *Journal of Interanational Development*, 2004, *16*, 1053–1066.

Arias, Omar, Gustavo Yamada, and Luis Tejerina, "Education, family background and racial earnings inequality in Brazil," *International Journal of Manpower*, 2004, *25* (3/4), 355–374.

Arnold, David and Joshua Bernstein, "The Effects of Tenure-Dependent Employment Protection Legislation," 2021. Unpublished manuscript.

Åslund, Olof, Lena Hensvik, and Oskar Nordström Skans, "Seeking Similarity: How Immigrants and Natives Manage in the Labor Market," *Journal of Labor Economics*, July 2014, *32* (3), 405–441.

Autor, David H. and David Scarborough, "Does Job Testing Harm Minority Workers? Evidence from Retail Establishments," *Quarterly Journal of Economics*, 2008, *123* (1), 219–277.

Azoulay, Pierre, Benjamin F. Jones, J. Daniel Kim, and Javier Miranda, "Age and High-Growth Entrepreneurship," *American Economic Review: Insights*, March 2020, *2* (1), 65–82.

Babina, Tania, "Destructive Creation at Work: How Financial Distress Spurs Entrepreneurship," *Review of Financial Studies*, September 2020, *33* (9), 4061–4101.

Bagger, Jesper and Rasmus Lentz, "An Empirical Model of Wage Dispersion with Sorting," *The Review of Economic Studies*, 2018, *86* (1), 153–190.

Barron, John M., Dan A. Black, and Mark A. Loewenstein, "Employer Size: The Implications for Search, Training, Capital Investment, Starting Wages, and Wage Growth," *Journal of Labor Economics*, January 1987, *5* (1), 76–89.

Bates, Timothy, "The Urban Development Potential of Black Owned Businesses," *Journal of the American Planning Association*, 2006, *72* (2), 227–237.

Bayer, Patrick, Stephen L. Ross, and Giorgio Topa, "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes," *Journal of Political Economy*, 2008, *116*, 1150–1194.

Becker, Gary S., *The Economics of Discrimination*, Unversity of Chicago Press, 1957.

Benson, Alan, Simon Board, and Moritz Meyer ter vehn, "Discrimination in Hiring: Evidence from Retail Sales," 2019. Unpublished manuscript.

Bernstein, Shai, Emanuele Colonnelli, Davide Malacrino, and Timtohy J. McQuade, "Who Creates New Firms when Local Opportunities Arise?," *Journal of Financial Economics*, 2021. Unpublished manuscript.

Bloom, Nicholas, Renata Lemos, Raffaella Sadun, Daniela Scur, and John Van Reenen, "JEEA-FBBVA Lecture 2013: The New Empirical Economics of Management," *Journal of the European Economic Association*, 2014, *12* (4), 835–876.

Bolte, Lukas, Nicole Immorlica, and Matthew O. Jackson, "The Role of Referrals in Immobility, Inequality, and Inefficiency in Labor Markets," November 2020. Unpublished manuscript.

Boston, Thomas D., "The Role of Black-Owned Businesses in Black Community Development," in Paul Ong and Anastasia Loukaitou-Sideris, eds., *Jobs and Economic Development in Minority Communities*, Temple University Press, 2006, chapter 7, pp. 161–175.

Brown, Charles and James Medoff, "The Employer Size-Wage Effect," *Journal of Political Economy*, October 1989, *97* (5), 1027–1059.

Brown, Meta, Elizabeth Setren, and Giorgio Topa, "Do Informal Referrals Lead to Better Matches? Evidence from a Firm's Employee Referral System," *Journal of Labor Economics*, January 2016, *34* (1), 161–209.

Buhai, I. Sebastian, Miguel A. Portela, Coen N. Teulings, and Aico van Vuuren, "Returns to Tenure or Seniority?," *Econometrica*, March 2014, *82* (2), 705–730.

Caldwell, Sydnee and Nikolaj Harmon, "Outside Options, Bargaining and Wages: Evidence from Coworker Networks," 2019. Unpublished manuscript.

Calvo-Armengol, Antoni and Matthew O. Jackson, "The Effects of Social Networks on Employment and Inequality," *American Economic Review*, 2004, *94*, 426–454.

Camara, M'Balou, Khaing Zaw, Darrick Hamilton, and William Darity Jr., "Entering Entrepreneurship: Racial Disparities in the Pathways into Business Ownership," Technical Report, The Samuel DuBois Cook Center on Social Equity October 2019.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline, "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

Cavalieri, Claudia and Reynaldo Fernandes, "Diferenciais de Salarios por Genero e por Cor: Uma Comparacao entre as Regioes Metropolitanas Brasileiras," *Revista de Economia Politica*, 1998, *18* (1), 158–175.

Cavounidis, Costas, Kevin Lang, and Russell Weinstein, "The Boss is Watching: How Monitoring Decisions Hurt Black Workers," February 2021. Unpublished manuscript.

Cingano, Federico and Alfonso Rosolia, "People I know: job search and social networks," *Journal of Labor Economics*, 2012, *30* (2), 291 – 332.

Cornell, Bradford and Ivo Welch, "Culture, Information, and Screening Discrimination," *Journal of Political Economy*, June 1996, *104* (4), 542–571.

Cornwell, Christopher, Ian M. Schmutte, and Daniela Scur, "Building a Productive Workforce: The Role of Structured Management Practices," *Management Science*, forthcoming, p. xxx.

_ , Jason Rivera, and Ian M. Schmutte, "Wage Discrimination When Identity is Subjective: Evidence from Changes in Employer-Reported Race," *Journal of Human Resources*, 2017, *52* (3), 719–755.

Correia, Sergio, Paulo Guimaraes, and Tom Zylkin, "PPMLHDFE: Fast Poisson Estimation with High-Dimensional Fixed Effects," *Stata Journal*, 2020, *20* (1), 95–115.

Daniel, G.R., *Race and Multiraciality in Brazil and the United States: Converging Paths?*, Pennsylvania State University Press, 2010.

Derenoncourt, Ellora, François Gérard, Lorenzo Lagos, and Claire Montialoux, "Racial Inequality, Minimum Wage Spillovers, and the Informal Sector," May 2021. Unpublished manuscript.

Dias, Marina and Roberto Hsu Rocha, "Racial Gaps in the Labor Market: The Role of Nonwhite Entrepreneurship," August 2021. Unpublished manuscript.

DiTomaso, Nancy, *The American Non-Dilemma: Racial Inequality Without Racism*, Russell Sage Foundation, 2013.

Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker, "Referral-based Job Search Networks," *Review of Economic Studies*, April 2016, *83* (2), 514–546.

Eliason, Marcus, Lena Hensvik, Francis Kramarz, and Oskar Nordström Skans, "Social Connections and the Sorting of Workers to Firms," 2020. Unpublished manuscript.

Fernandez, Roberto M. and Isabel Fernandez-Mateo, "Networks, Race, and Hiring," *American Sociological Review*, February 2006, *71* (1), 42–71.

_ **and M. Lourdes Sosa**, "Gendering the Job: Networks and Recruitment at a Call Center," *American Journal of Sociology*, 2005, *111* (3), 859–904.

_ **, Emilio J. Castilla, and Paul Moore**, "Social Capital at Work : Networks and Employment at a Phone Center," *American Journal of Sociology*, 2000, *105* (5), 1288–1356.

Fiola, J., *Race Relations in Brazil: A Reassessment of the "Racial Democracy" Thesis* Occasional papers series, Latin American Studies Program. University of Massachusetts at Amherst, 1990.

Fisman, Raymond, Daniel Paravisini, and Vikrant Vig, "Cultural Proximity and Loan Outcomes," *American Economic Review*, February 2017, *107* (2), 457–492.

Francis, Andrew M. and Maria Tannuri-Pianto, "Endogenous Race in Brazil : Affirmative Action and the Construction of Racial Identity among Young Adults," *Economic Development and Cultural Change*, 2013, pp. 1–35.

Galenianos, Manolis, "Learning about match quality and the use of referrals," *Review of Economic Dynamics*, October 2013, *16* (4), 668–690.

Garia, Luana Marques, Hugo Nopo, and Paola Salardi, "Gender and Racial Wage Gaps in Brazil 1996-2006: Evidence Using a Matching Comparisons Approach," 2009. Inter-American Development Bank Research Department Working Paper No. 681.

Gee, Laura, Jason Jones, and Moira Burke, "Social Networks and Labor Markets: How Strong Ties Relate to Job Finding on Facebook's Social Network," *Journal of Labor Economics*, 2017, *35* (2), 485 – 518.

Gee, Laura K., Jason J. Jones, Christopher J. Fariss, Moira Burke, and James H. Fowler, "The paradox of weak ties in 55 countries," *Journal of Economic Behavior I& Organization*, 2017, *133*, 362–372.

Gerard, Francois, Lorenzo Lagos, Edson Severnini, and David Card, "Assortative Matching or Exclusionary Hiring? The Impact of Employment and Pay Policies on Racial Wage Differences in Brazil," *American Economic Review*, forthcoming.

Giuliano, Laura, David I. Levine, and Jonathan Leonard, "Racial Bias in the Manager-Employee Relationship: An Analysis of Quits, Dismissals, and Promotions at a Large Retal Firm," *Journal of Human Resources*, January 2011, *46* (1), 26–52.

Glitz, Albrecht, "Coworker networks in the labour market," *Labour Economics*, 2017, *44*, 218–230.

Haanwinckel, Daniel and Rodrigo R. Soares, "Workforce Composition, Productivity, and

Labor Regulations in a Compensating Differentials Theory of Informality," Working Papers 2020-45, Becker Friedman Institute for Research In Economics April 2020.

Hellerstein, Judith K., Mark J. Kutzbach, and David Neumark, "Do labor market networks have an important spatial dimension?," *Journal of Urban Economics*, 2014, *79*, 39–58.

_ , Melissa McInerney, and David Neumark, "Neighbors and Co-Workers: the Importance of Residential Labor Market Networks," *Journal of Labor Economics*, 2011, *29* (4), 659–695.

Hensvik, Lena and Oskar Nordström Skans, "Social Networks, Employee Selection, and Labor Market Outcomes," *Journal of Labor Economics*, October 2016, *34* (4), 825–867.

Hirata, Guilherme and Rodrigo R. Soares, "Competition and the Racial Wage Gap: Evidence from Brazil," *Journal of Development Economics*, 2020, *146.*

Holzer, Harry J., "Hiring Procedures in the Firm: Their Economic Determinants and Outcomes," March 1987. NBER Working Paper 2185.

_ , "Informal Job Search and Black Youth Unemployment," *American Economic Review*, June 1987, *77* (3), 446–452.

_ , "Why Do Small Establishments Hire Fewer Blacks Than Large Ones?," *Journal of Human Resources*, 1998, *33* (4), 896–914.

_ , Steven Raphael, and Michael A. Stoll, "Perceived Criminality, Criminal Background Checks, and the Racial Hiring Practices of Employers," *Journal of Law and Economics*, 2006, *49* (2), 451–480.

Kerr, Sari Pekkala and William R. Kerr, "Immigrant Entrepreneuship," in John Haltiwanger, Erik Hurst, Javier Miranda, and Antoinette Schoar, eds., *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*, University of Chicago Press, 2017.

_ and _ , "Whose Job Is It Anyway? Co-Ethnic Hiring in New U.S. Ventures," February 2021. NBER Working Paper 28509.

Kramarz, Francis and Oskar Nordström Skans, "When Strong Ties are Strong: Networks and Youth Labour Market Entry," *The Review of Economic Studies*, 2014.

Lang, Kevin, "A Language Theory of Discrimination," *Quarterly Journal of Economics*, May 1986, *101* (2), 363–382.

_ and Jee-Yeon K. Lehmann, "Racial Discrimination in the Labor Market: Theory and Empirics," *Journal of Economic Literature*, 2012, *50* (4), 959–1006.

Lehmann, Jee-Yeon K., "Job Assignment and Promotion Under Statistical Discrimination: Evidence from the Early Careers of Lawyers," October 2013. Unpublished manuscript.

Lovell, Peggy A., "Race, Gender, and Development in Brazil," *Latin American Research Review*, 1994, *29* (3), 7–35.

Lundberg, Shelly J. and Ricard Startz, "Private Discrimination and Social Intervention in Competitive Labor Market," *American Economic Review*, June 1983, *73* (3), 340–347.

Machado, Marta, Márcia Regina Barros da Silva, and Natália Neris da Silva Santos, "Anti-racism legislation in Brazil: the role of the Courts in the reproduction of the myth of racial democracy," *Revista de Investigações Constitucionais*, 2019, *6*, 267–296.

Marsden, Peter V., "The hiring process," *American Behavioral Scientist*, 1994, *37* (7), 979–991.

Matos, Raquel S. and Ana F. Machado, "Diferencial de Rendimentos por Cor e Sexo no Brasil (1987-2001)," *Economica*, 2006, *8* (1), 5–27.

McPherson, Miller, Lynn Smith Lovin, and James M. Cook, "Birds of Feather: Homophily in Social Networks," *Annual Review of Sociology*, 2001, *27*, 415–444.

Miller, Conrad, "The Persistent Effect of Temporary Affirmative Action," *American Economic Journal: Applied Economics*, July 2017, *9* (3), 152–190.

Morgan, John and Felix Várdy, "Diversity in the Workplace," *American Economic Review*, March 2009, *99* (1), 472–485.

Okafor, Chika O., "All Things Equal: Social Networks as a Mechanism for Discrimination," December 2020. Unpublished manuscript.

Pedulla, David S. and Devah Pager, "Race and Networks in the Job Search Process," *American Sociological Review*, 2019, *84* (6), 983–1012.

Petersen, Trond, Ishak Saporta, and Marc-David L. Seidel, "Offering a Job: Meritocracy and Social Networks," *American Journal of Sociology*, November 2000, *106* (3), 763–816.

Rebien, Martina, Michael Stops, and Anna Zaharieva, "Formal Search and Referrals From a Firm's Perspective," *International Economic Review*, November 2020, *61* (4), 1679–1748.

Reis, Mauricio Cortez and Ana Risi Vianna Crespo, "Race Discrimination in Brazil: An Analysis of the Age, Period and Cohort Effects," 2015. Institute for Applied Economic Research (IPEA) Discussion Paper No. 1114.

Restuccia, Diego and Richard Rogerson, "The Causes and Costs of Misallocation," *Journal of Economic Perspectives*, Summer 2017, *31* (3), 151–174.

San, Shmuel, "Who Works Where and Why? Parental Networks and the Labor Market," May 2021. Unpublished manuscript.

Saygin, Perihan Ozge, Andrea Weber, and Michèle A. Weynandt, "Coworkers, Networks, and Job-Search Outcomes among Displaced Workers," *ILR Review*, 2021, *74* (1), 95–130.

Schmutte, Ian M., "Job Referral Networks and the Determination of Earnings in Local Labor Markets," *Journal of Labor Economics*, 2015, *33* (1), 1–32.

Silva, Nelson V., "O Preco da Cor: Diferenciais Raciais na Distribuicao da Renda no Brasil," *Pesquisa e Planejamento Economico*, 1980, *10* (1), 21–44.

_ , "Updating the Cost of Not Being White in Brazil," in P.M. Fontaine, ed., *Races, Classes and Power in Brazil*, Los Angeles: UCLA Press, 1985, pp. 42–55.

Simon, Curtis J. and John T. Warner, "Matchmaker, Matchmaker: The Effect of Old Boy Networks on Job Match Quality, Earnings, and Tenure," *Journal of Labor Economics*, July 1992, *10* (3), 306–330.

Small, Mario L. and Devah Pager, "Sociological Perspectives on Racial Discrimination," *Journal of Economic Perspectives*, Spring 2020, *34* (2), 49–67.

Sorkin, Isaac, "Ranking Firms Using Revealed Preference," *The Quarterly Journal of Economics*, 2018, *133* (3), 1331–1393.

Telles, Edward E., *Race in Another America: The Significance of Skin Color in Brazil*, Princeton, NJ: Princeton University Press, 2004.

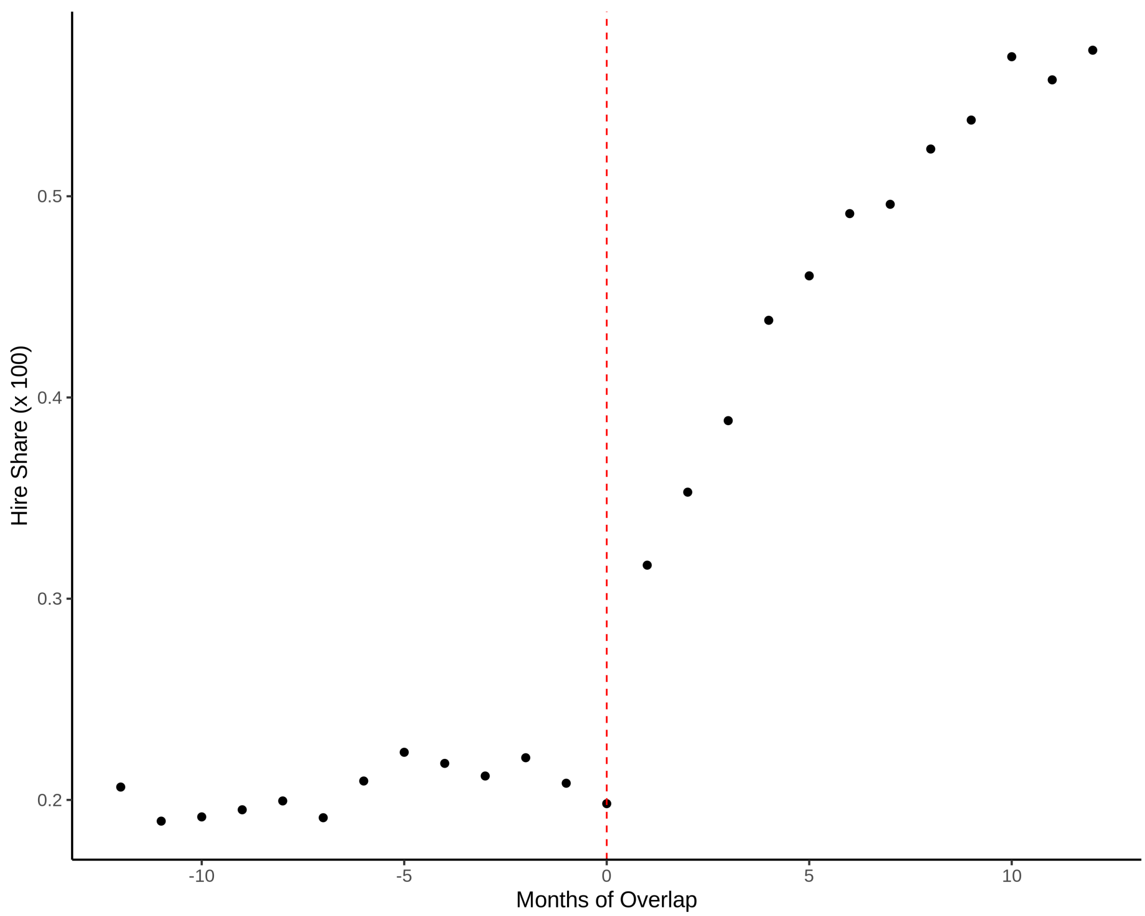
Topa, Giorgio, "Social Interactions, Local Spillovers and Unemployment," *Review of Economic Studies*, 2001, *68*, 261–295.

_ , "Social and spatial networks in labour markets," *Oxford Review of Economic Policy*, 2019, *35* (4), 722–745.

Wooldridge, Jeffrey M., "Distribution-free estimation of some nonlinear panel data models," *Journal of Econometrics*, 1999, *90* (1), 77–97.

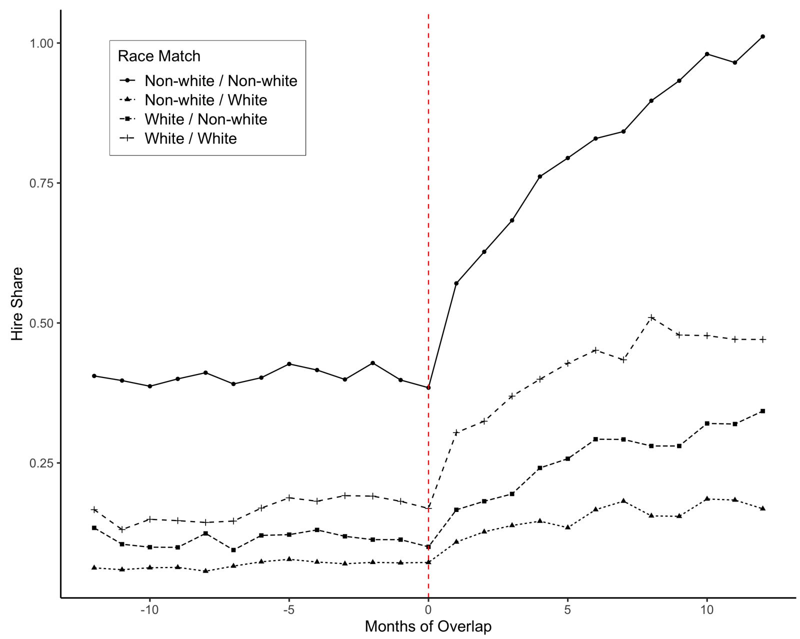
Wozniak, Abigail, "Discrimination and the Effects of Drug Testing on Black Employment," *Review of Economics and Statistics*, 2015, *97* (3), 548–566.

Figure 1
True Coworker Connections Relative to Placebo Connections



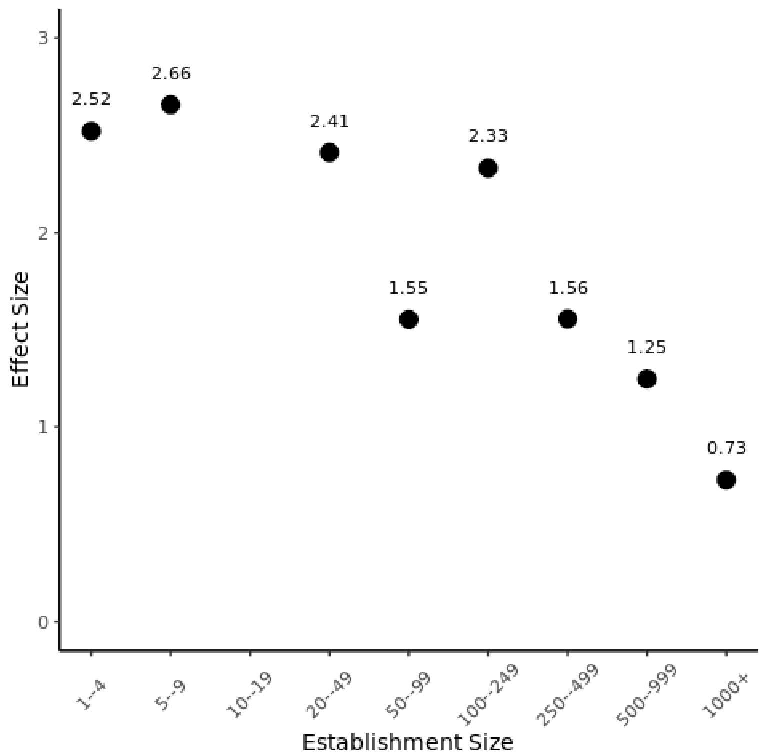
Note: This figure plots the average value of the hiring outcome, P_{ijk} from in equation (1) for observations where there is overlap in the previous employment history of job seeker i and an incumbent worker at establishment k. The horizontal axis measures the number of months that two workers overlapped in a previous job. When overlap is negative, it measures the number of months that passed between the two workers' spells. Note that the hiring outcome is scaled up by 100 to reflect percentage point changes.

Figure 2
Hiring Share by Coworker Overlap by Incumbent and Job Seeker Race



Note: This figure plots the average value of the hiring outcome in equation (1) relative to job spell overlap. The horizontal axis measures *overlap*, the number of months that two workers overlapped in a previous job. When overlap is positive, the two workers were true coworkers for that number of months. When it is negative, it measures the number of months that passed between the two workers' spells. When reporting the race match, we put the race of the job seeker first and the linked incumbent second. So "White / Nonwhite" indicates a white job seeker is linked to a nonwhite incumbent at the destination.

Figure 3
Referral Effects Decreasing in Establishment Size



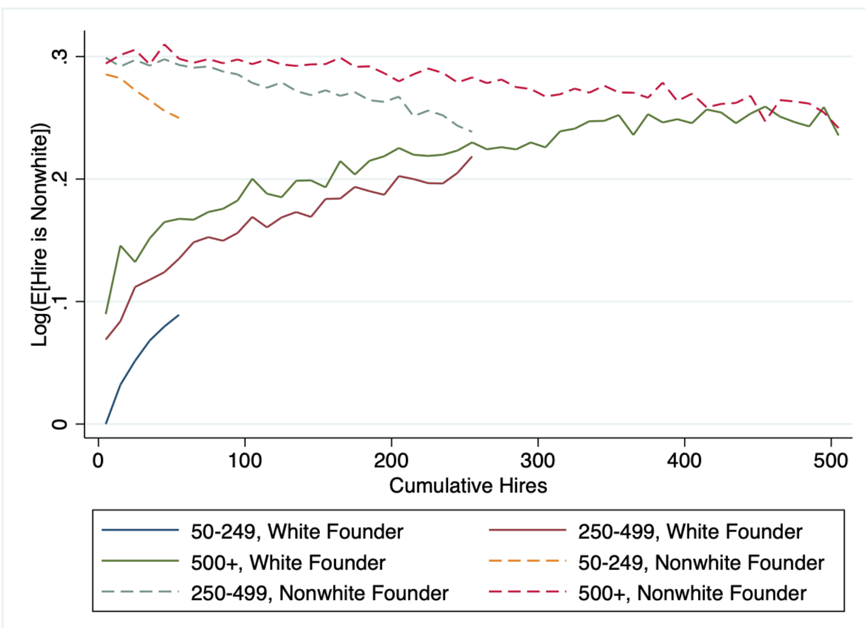
Note: This figure plots point estimates of referral effects by destination establishment size from estimating equation (21). The plotted coefficients represent the effect of a true link on hiring as a share of the average hiring rate in dyads involving destinations in the reported size class.

Figure 4

Nonwhite Share of Hires Converges with Cumulative Hires

(a) Pooled

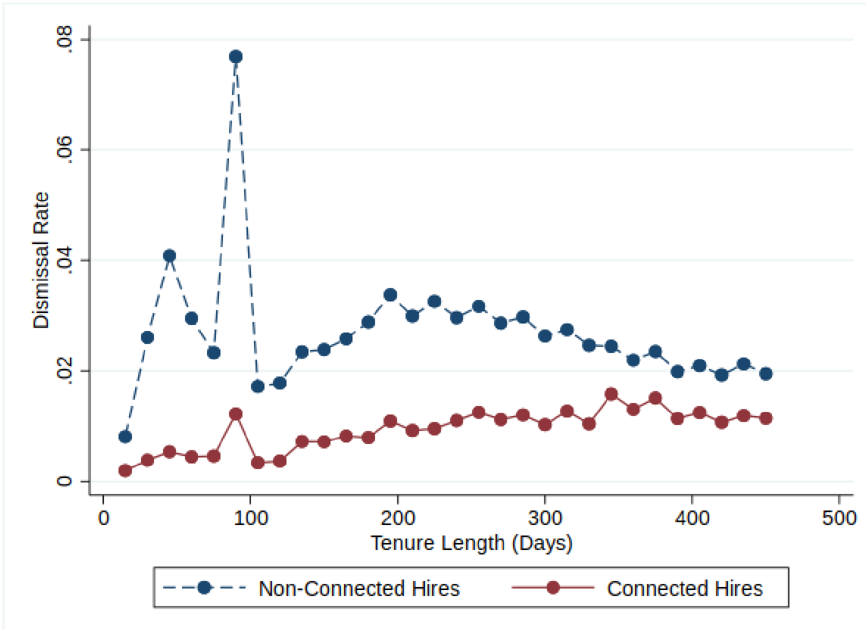
(b) By Total Hires, Balanced



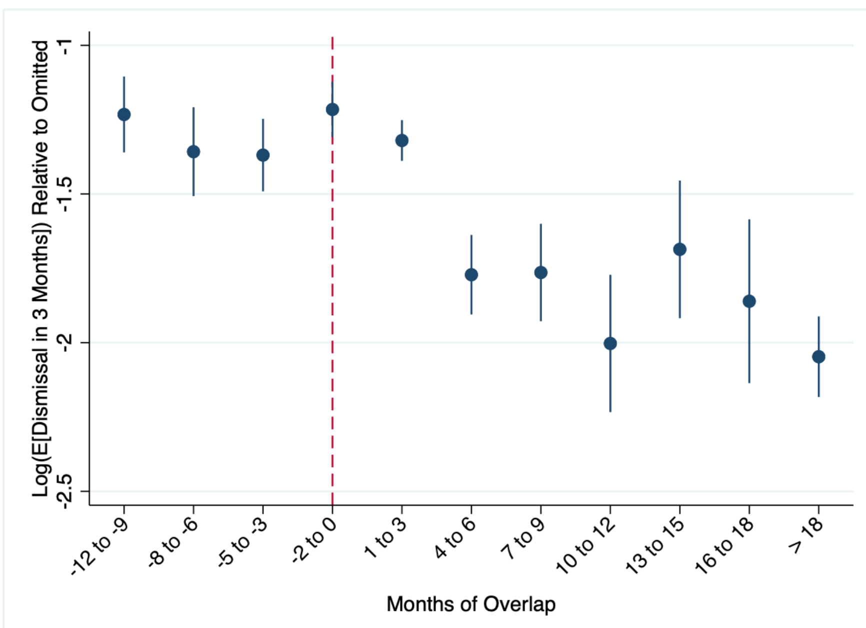
Note: This figure plots the relationship between the racial composition of a firm's hires and its cumulative hires to date. Panel A plots the $\eta^{n,r}$ coefficient estimates from equation (3), summarizing the relationship between a firm's racial composition of hires, its cumulative hires to date (n), and the race of its founder (r). Panel B plots the $\eta^{s,n,r}$ coefficient estimates from equation (4), which allows the relationship between a firm's racial composition of hires, its cumulative hires to date, and the race of the founder to vary with the firm's total observed hires (s). Both models are estimated via Poisson quasi maximum likelihood (PQML). In Panel A the omitted category is the first five hires after the year of entry for firms with white founders. In Panel B the omitted category is the first ten hires after the year of entry for firms with white founders and 50-249 total observed hires. Founder race is inferred from the race of the top-paid manager or employee at entry.

Figure 5
Connected Hires Are Less Likely to Be Dismissed

(a) Dismissal Rates by Job Tenure and Connected Status

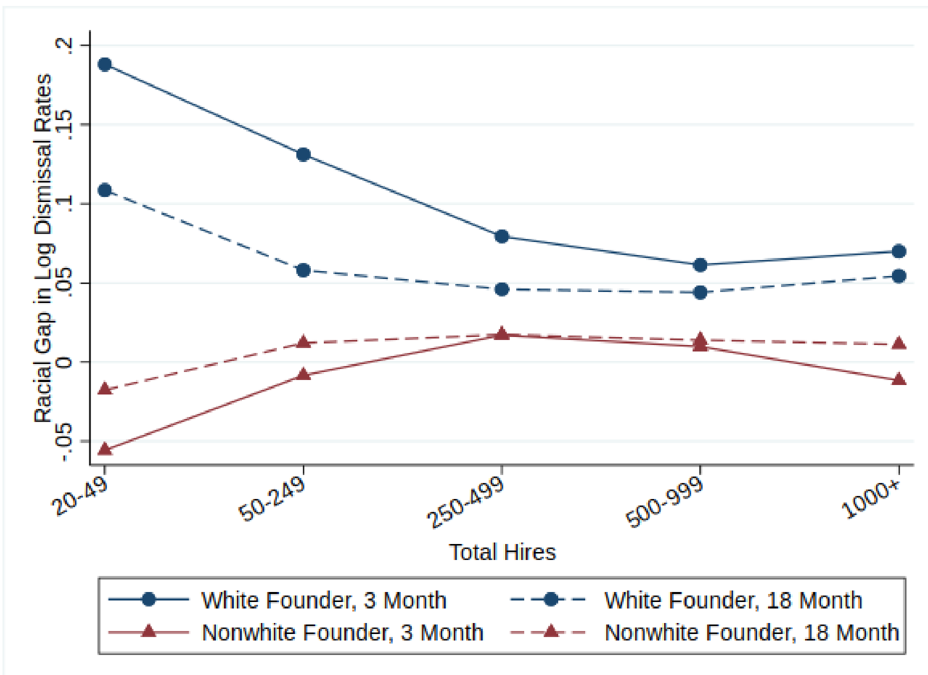


(b) 3-Month Dismissal Rate by Overlap



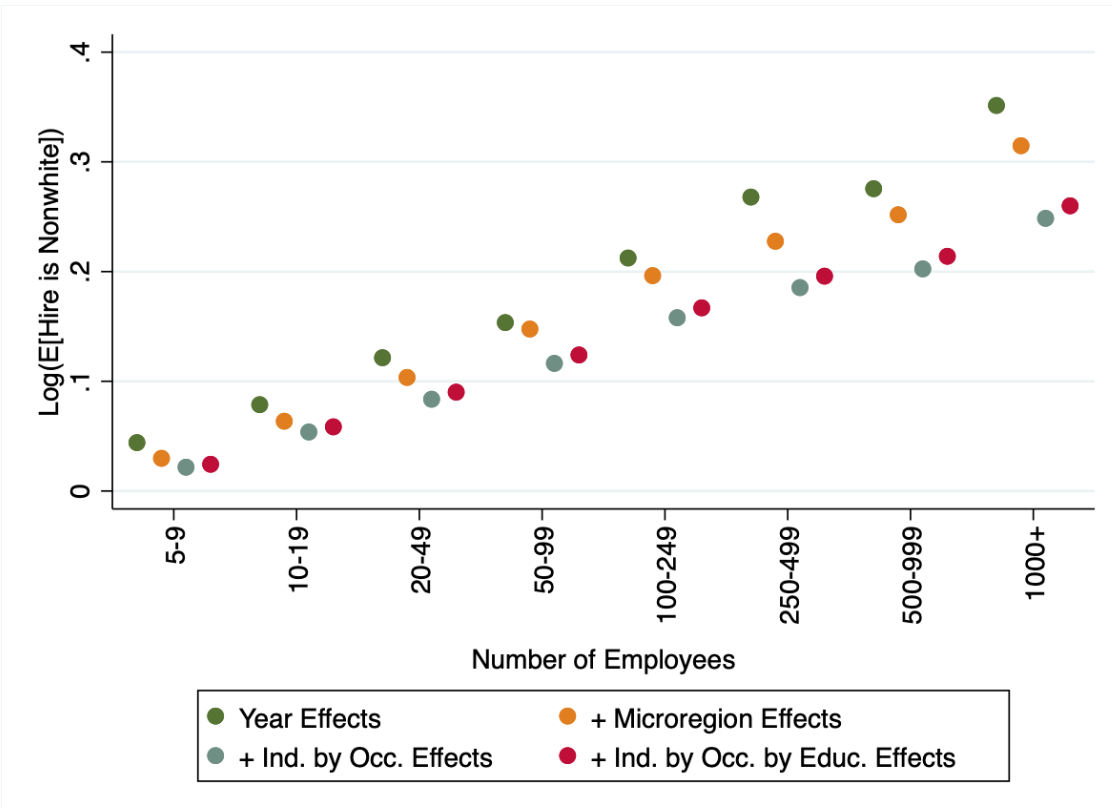
Note: Panel A plots dismissal rates by job spell tenure for connected and non-connected hires, adjusting for establishment, occupation, and year fixed effects as described in equation (5). The model is estimated via Poisson quasi maximum likelihood (PQML). "Connected" hires had previously shared a workplace (with no more than 100 employees) with an incumbent at this establishment.

Figure 6
Racial Disparity in Dismissal Rates by Total Hires



Note: This figure plots the adjusted, firm-level nonwhite-white difference in log dismissal rates as a function of founder race and the firm's total number of observed hires after the year of entry. Firm-specific racial differences in dismissal rates are constructed as described in equation (7). The model is estimated via Poisson quasi maximum likelihood (PQML). We limit to establishments with 20 or more hires.

FIGURE 7
NONWHITE SHARE OF NEW HIRES BY ESTABLISHMENT SIZE



Note: This figure plots $\sigma_{c(J(i,t))}$ coefficient estimates for several specifications of the model (13) described in Section 6.3, which are estimated via Poisson quasi maximum likelihood (PQML). *Year Effects* refers to a model that includes only year effects (τ_t) as additional controls. *+ Microregion Effects* refers to a model that includes microregion fixed effects ($\mu_{m(J(i,t))}$) in addition to year effects. *+ Ind. by Occ. Effects* refers to a model that also includes fixed effects for three-digit occupation by two-digit industry combinations. *+ Ind. by Occ. by Educ. Effects* replaces industry by occupation fixed effects with fixed effects for industry by occupation by worker education category fixed effects.

Table 1
Entrepreneurship Rates and Characteristics of Private Sector Employees by Race Group

	All (1)	White (2)	Mixed (3)	Black (4)
A: Men				
Share of sample in column race group	1.00	0.48	0.43	0.08
Share in private employment	0.39	0.41	0.37	0.42
Share unemployed	0.051	0.045	0.056	0.064
Share entrepreneurs	0.030	0.041	0.021	0.018
Characteristics of private sector employees				
Mean years of education	8.67	9.40	7.91	7.96
Fraction with HS or more	0.47	0.54	0.39	0.39
Mean log hourly wage	1.96	2.08	1.81	1.92
Share in formal sector employment	0.76	0.79	0.72	0.76
A: Women				
Share of sample in column race group	1.00	0.50	0.42	0.08
Share in private employment	0.21	0.24	0.17	0.19
Share unemployed	0.065	0.056	0.071	0.086
Share entrepreneurs	0.016	0.022	0.010	0.008
Characteristics of private sector employees				
Mean years of education	10.33	10.77	9.70	9.62
Fraction with HS or more	0.68	0.72	0.62	0.62
Mean log hourly wage	1.88	1.97	1.74	1.85
Share in formal sector employment	0.78	0.80	0.74	0.77

This table reports statistics from the Pesquisa Nacional por Amostra de Domicilios (PNAD) household survey for the years 2003 through 2015. The sample is limited to men and women ages 18 to 65. We define entrepreneurs as those who self-report running a business, formal or informal with at least one paid employee.

TABLE 2

RAW DIFFERENCES IN WORKER AND JOB CHARACTERISTICS BY RACE

	All Employees			Recent Hires					
				All Firms			Entrant Firms		
	Pooled (1)	White (2)	Nonwhite (3)	Pooled (4)	White (5)	Nonwhite (6)	Pooled (7)	White (8)	Nonwhite (9)
Nonwhite (%)	36.5	0.0	100.0	39.0	0.0	100.0	36.0	0.0	100.0
Log Wage	2.006	2.079	1.878	1.851	1.900	1.775	1.840	1.877	1.777
	(0.674)	(0.713)	(0.581)	(0.553)	(0.582)	(0.493)	(0.455)	(0.468)	(0.421)
Male (%)	66.3	64.3	69.9	67.4	64.8	71.4	65.5	63.4	69.3
Age	33.7	34.0	33.1	30.9	31.1	30.7	31.0	31.2	30.7
< HS	30.4	28.7	33.5	28.4	25.9	32.3	21.5	19.6	24.8
HS Grad	57.2	56.3	58.8	61.7	61.8	61.4	69.5	69.7	69.1
College Grad	12.4	15.0	7.7	9.9	12.2	6.3	9.1	10.7	6.0
Number of Worker-Year Obs.	688m	437m	251m	254m	155m	99m	55m	35m	20m

This table reports summary statistics from the *Relação Anual de Informações Sociais* (RAIS) data for the years 2003-2017. We limit the sample to the jobs of men and women between the ages of 18 and 65 on private sector, indeterminate-length contracts for at least 30 hours per week. Columns 1–3 report statistics for all job spell-years. Columns 4–9 report statistics for the first year of a job spell. Columns 7-9 restrict to entrant firms as described in Section 4.2.

Table 3
Descriptive Statistics for Referral Analysis Sample

	Dyads (1)	Job Changer (2)	Incumbents (3)
Any Link	8.4%		
Linked	4.1%		
Hired	0.082%		
White	30.2%	32.0%	50.1%
Male	43.2%	55.9%	62.4%
Age	32.2	31.7	34.2
Dest. Size			
1–99	62.6%	59.0%	60.5%
100–499	21.0%	22.7%	20.2%
500+	16.5%	18.3%	19.3%
Num. Obs.	303,338,866	1,353,787	9,216,640

Source: RAIS, 2013–2017. The "Dyads" column includes pairs of job changers matched to potential destinations. The "Incuments" column describes the population of incumbent workers who are linked to some hired worker via a past coworking relationship.

Table 4
Referral Effects by Job Seeker and Incumbent Race

	Overall			Race Match
	(1)	(2)	(3)	(4)
True Link	0.182	0.222	0.117	
	(0.003)	(0.004)	(0.003)	
Any Link		0.097	0.066	
		(0.004)	(0.002)	
Race Match × True Link				
Nonwhite / Nonwhite				0.192
				(0.007)
Nonwhite / White				0.025
				(0.004)
White / Nonwhite				0.020
				(0.007)
White / White				0.136
				(0.006)
Dep. Var. Mean.	0.084	0.084	0.084	0.084
Estab. Pair FE	✓		✓	✓
Placebo Link Control		✓	✓	✓
Num. Estab. Pairs		23,026,153		
Number of Obs.		303,338,866		

Columns 1–3 presents estimated referral effects under different identifying assumptions. Column 4 reports heterogeneity in referral effects based on the match between the race of the job changer and the race of the linked incumbent. All specifications include controls for worker demographic and human capital characteristics. Column 2 controls for origin and destination establishment effects. Column 4 includes controls for each race match interacted with "Any Link", which indicates observations for which the job changer has either a true coworker or a placebo coworker connection to an incumbent worker at the destination. When reporting the race match, we put the race of the job seeker first and the linked incumbent second. So "White / Nonwhite" indicates a white job seeker is linked to a nonwhite incumbent at the destination.

TABLE 5

RACIAL DIFFERENCES IN DISMISSAL RATES AND JOB SENIORITY, BY FOUNDER RACE

	All Entrants			White Founders			Nonwhite Founders		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A									
Outcome: Dismissal During Probation									
Nonwhite	0.083	0.042	0.075	0.121	0.081	0.114	0.042	0.011	0.036
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓
Occupation FEs		✓			✓			✓	
Education FEs			✓			✓			✓
Number of Obs.		52,376,661			35,138,895			17,237,766	
Panel B									
Outcome: Seniority Index									
Nonwhite	-0.049	-0.040	-0.050	-0.097	-0.086	-0.099	0.002	0.005	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓
Occupation FEs		✓			✓			✓	
Education FEs			✓			✓			✓
Number of Obs.		39,203,654			26,171,760			13,031,894	

This table presents regression coefficient estimates for equation (10) in Panel A and equation (12) in Panel B. In Panel A the outcome is an indicator for whether a job spell results in a dismissal within three months of the hire date. Each observation is a job spell. In Panel B the outcome is the *seniority index*, which summarizes an employee's tenure relative to their colleagues, and is defined in equation (11). Each observation is a job spell-year. Columns 1 through 3 pool all entrant firms, columns 4 through 6 limits to firms with white founders, and columns 7 through 9 limit to firms with nonwhite founders.

Online Appendix: The Dynamics of Referral Hiring and Racial Inequality: Evidence from Brazil

Conrad Miller

Ian Schmutte

August 2021

A Data Appendix

A.1 RAIS Data

We prepare the RAIS data in several steps. First, we clean the raw data files retrieved from the MTE. Next, we prepare a master dataset that imposes certain variable definition and data cleaning decisions. Finally, we prepare the various samples that are needed for particular analyses.

A.1.1 Cleaning the Raw Data

The raw data files are delivered by year, and our analysis in this paper uses the data from 2003–2017. The variables available change across years, as does their coding. In a first step, we build a codebook and redefine variable names and labels to better track relationships among the variables.

Workers are uniquely identified by a PIS code and establishments by a CNPJ code. We build a relational database comprised of four tables:

- `Job` table with a single record for each PIS-CNPJ-YEAR that includes all characteristics specific to the employment match.
- `Establishment` table with a single record for each CNPJ-YEAR pair with all characteristics specific to an establishment.
- `Worker` table, with a single record for each PIS.

To prepare the `Job` table, we first disambiguate a handful of records that duplicate the same PIS-CNPJ pair in the same year. In a small fraction (less than 2 percent) of cases, the raw data have multiple records for the same PIS-CNPJ pair in a given year. A negligible number (around 15 per year out of roughly 60 million) also share the same reported date of hire. The vast majority (95-98 percent) are pairs with exactly 2 records in the same year. The extra records are associated with administrative reassignments that are not consequential for our analysis, and mostly occur in public-sector jobs. In all cases, we combine the repeated records into a single PIS-CNPJ-year level record that includes all earnings information, the earliest date of hire, and all other characteristics from the record with the latest date of separation. After completing this disambiguation, each record is uniquely identified by a combination of PIS-CNPJ-YEAR. For variables whose coding changes over time (like education and race), we define a harmonized version that has a consistent coding across all years.

To prepare the `Establishment` table, we compute the modal value for each establishment characteristic (industry, size class, location, ownership type) across all job-level records in the `Job` table.

An important feature of the RAIS data is that establishments can, and do, report different values for the demographic characteristics of the same PIS (Cornwell et al., 2017). The `Worker` table includes the modal values for race, gender, and date of birth across all records in the `Jobs` table that involve the same PIS. We also retain the time-varying information on employer-reported race, gender, date of birth, and education in the `Job` table. We also define an additional measure of education which records, for each year, the highest level of education reported for that PIS up to that date.

A.1.2 Primary Analysis Data

From the cleaned database, we extract primary analysis data for each of Brazil's five regions. We impose very few restrictions at this stage, but define a few key variables:

Wages: the real hourly wage (in 2015 Brazilian Reais). We divide real monthly earnings by the number of contracted hours per month. To approximate the number of hours a worker is contracted to work each month, we multiply contracted hours per week, which is reported in RAIS, by $\frac{30}{7}$. Average monthly earnings are reported in nominal reais, which we convert to constant 2015 reais using the OECD's Consumer Price Index for Brazil.

Dominant Job: In much of the literature, and our analysis, it is common to assemble a worker-year panel from the linked data. Since workers often hold multiple jobs in the same year, we define the *dominant job* as the job with highest earnings for the year among all those with the longest observed tenure.

Valid Identifiers: The PIS and CNPJ numbers are social security and tax identifiers that include check digits, by which it is possible to identify records with invalid identifiers.

A.1.3 Data for Referral Analysis

To study referrals, we first extract data on new hires from the primary analysis data. We restrict the sample to job-year observations with valid PIS and CNPJ identifiers. The RAIS data provide several different ways to identify new hires, and we require they all agree. Specifically, we extract PIS-CNPJ-YEAR observations when they are (1) the first time a PIS-CNPJ combination is observed; (2) the match is actually coded as a new hire; (3) the recorded year of hire corresponds to the year of the observation. For each new hire we link information on their prior year employer, including those who were not employed in the prior year. For our analysis of displaced workers, we restrict the sample to those newly-hired workers whose prior-year employer had a mass displacement event.[33]

To define coworking relationships, we extract all job-year observations with valid CNPJ and PIS data for 2003–2015, keeping only full-time jobs (at least 35 hours contracted per week) and in establishments with at least 4 and fewer than 100 employees. We then define a dataset with one observation for each PIS-CNPJ employment match that records the start and end dates of the job spell. Next, we form the full cartesian product of the PIS-CNPJ level data, joined by CNPJ, which forms one observation for each pair of workers ever employed in the same CNPJ. For each such pair, we compute the overlap in their job spells as the number of months between the start of the later-starting job and the end of the earlier-ending job. If the earlier-ending job ends after the later-starting job starts, overlap is positive. Otherwise, it is negative. We retain all pairs with overlap greater than or equal to -12 months.

To build the dyad data, we associate each worker hired in a given year t with the plant from which they separated in year $t-1$ for all hires between 2013–2017. For each origin firm, we restrict attention to potential destinations to which at least on separating worker from the origin plant moves. For each separating worker, we assign one observation for each such potential destination. Then, using the information on overlapping coworker pairs, we define "linked" potential destinations as those where the separating worker has overlap at least one incumbent worker. Finally, we link basic demographic information for the focal (hired) worker and for the linked incumbent (when there is one).

To ensure that our analysis of referrals is not simply picking up tied moves where multiple workers from the same plant all move to the same destination, we do several things. First, we ensure that the linked incumbent in the potential destination was not hired in the same year the separating worker is at risk to move there. Second, we only use coworking relationships that were formed at least two years prior to the move. Finally, we make sure that the plant at which the two workers were most recently employed together was neither the origin firm for the separating worker, nor the potential destination where the linked incumbent is employed.

A.1.4 Estimation of the AKM model

For certain analyses, we use employer effects from the canonical two-way fixed effects model introduced by Abowd et al. (1999), which models the log wage as a linear function of unobserved worker and employer heterogeneity. As is standard, we estimate the model using the pre-conditioned conjugate gradient algorithm (`pcg` in MATLAB) and then separately identify the firm and worker effects within each connected component of the realized mobility network. See Abowd et al. (2002) for details regarding the estimation and identification methods.

We estimate the AKM model separately by region, restricting the sample to dominant job contract-years where: both the PIS and CNPJ are valid, average monthly earnings are positive, and the employed worker is between 20 and 60 years of age. We control for time-varying worker

[33] We define mass displacement events as those where the establishment's employment drops by between 60 and 90 percent in a single year.

characteristics: a cubic in age interacted with race, gender, and education, along with a full set of unrestricted year effects. To ensure the worker effects are separately identified relative to the year effects and linear term in age, we normalize the age profile to flatten out at age 30 (Card et al., 2018).

B Appendix: Additional Results

B.1 Racial Differences in Preferences over Employer Size

One alternative explanation for the sorting pattern we document is that, relative to white job seekers, nonwhite job seekers have a strong preference for working at large employers. We build on the insight that, under some assumptions, worker preferences over employers can be inferred from worker mobility patterns (e.g. Sorkin, 2018). We use the *poaching rank* as developed by Bagger and Lentz (2018). The premise of the poaching rank is that higher-ranked employers should hire relatively more workers from employment than from unemployment. That's because poaching a worker from another employer indicates that the worker prefers the destination employer. The poaching index for establishment J is defined as the share of all new hires that are poached from other employers:

$$p_J = \frac{n(.,J)}{n(0,J) + n(.,J)} \tag{B.1}$$

where $n(.,J)$ is number of hires poached from other and $n(0,J)$ hires from unemployment.

The poaching rank of establishment J is simply a conversion of the poaching index into an ordinal ranking of employers. We group establishments into 1000 quantiles based on their poaching index.

We measure race-specific poaching ranks for each establishment, and then relate those ranks to founder race and total hires using the sample of entrants discussed in Section 4.2. Figure B.12 reports the results.

B.2 Analysis of Displaced Workers

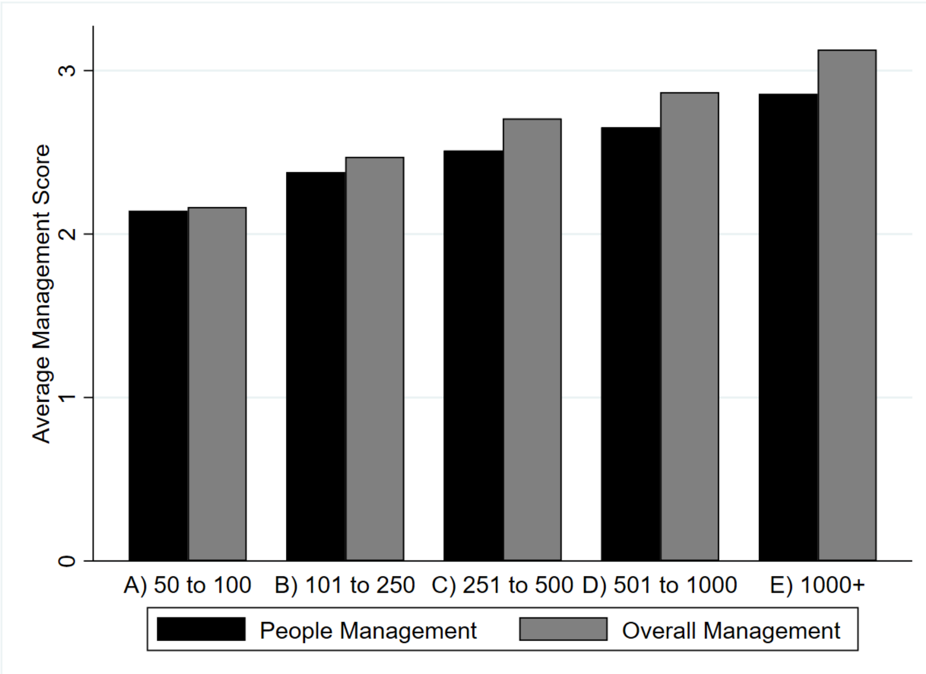
Our main estimates of referral effects in Section 4.1.2 are based on data covering all workers that separated from jobs between 2012 and 2016. Because these workers may have been inspired to change jobs due to the quality of their social networks, our results could partially be driven by self-selection. To address this concern, we have re-estimated referral effects focusing only on workers that were displaced from their employers. First, we identify mass displacement events in the RAIS data as those years in which plant-level employment contracts by between 60 and 90 percent from a baseline of at least ten employees.[34] Then we construct dyads for those workers that separate from plants where there was a mass displacement event, just as in Eliason et al. (2020). The rest of the data construction is identical to the sample of all job changers.

[34] We exclude events in which plant employment goes to zero to avoid capturing plant acquisitions or mergers in our data.

Table B.1 reports descriptive statistics for the displaced worker sample. Of note, the share of dyads that record a hire is smaller (0.074 percent) than the full sample (0.082 percent). The number of dyads that in which the displaced worker has a coworker link to the target firm is also smaller (3.7 percent versus 4.1 percent in the full sample). There are just 39,701 displaced workers in our sample. They are slightly older, less likely to be white, and considerably more likely to be male, than the 1.4 million workers in the full data.

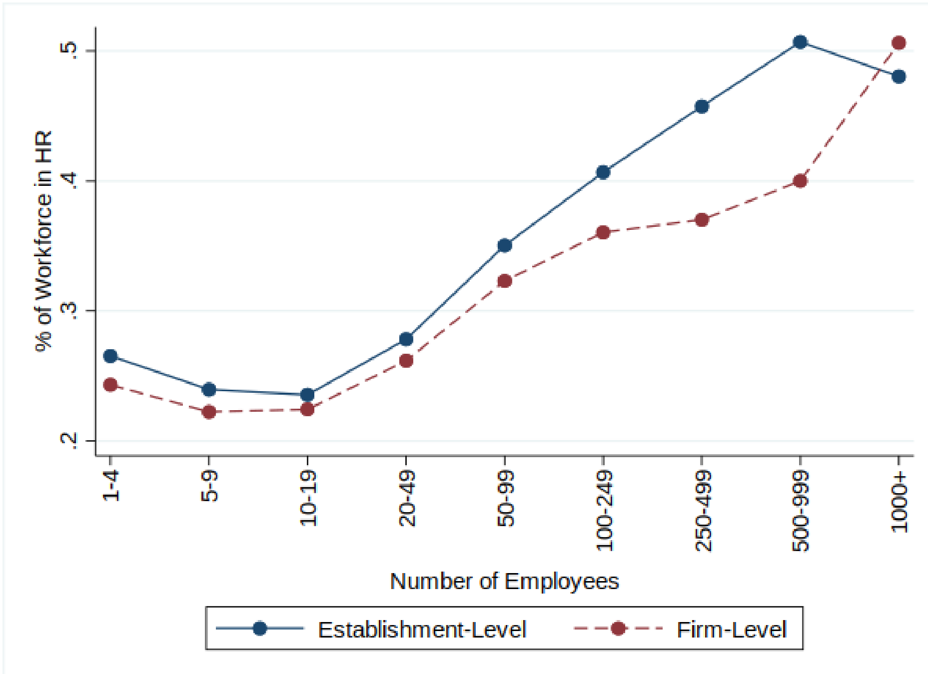
Table B.2 reports the same models as Table 4 for the displaced worker sample. The results for the displaced workers sample are nearly identical to the results based on the full sample. The point estimates in our preferred specifications (Columns 3 and 4) are slightly smaller, though given differences in the baseline mean and the precision of the estimates, one would be hard pressed to make a strong claim that the quantitative differences are meaningful. We conclude that bias driven by selection of job movers is not driving our main results.

Figure B.1
World Management Survey Scores Increasing in Employer Size



Note: This figure reports average overall management and people management scores for medium and large Brazilian manufacturing firms in the World Management Survey. The scores measure the adoption of formal management practices in different areas of performance, including personnel (people) management, operations, and target-setting. See Bloom et al. (2014).

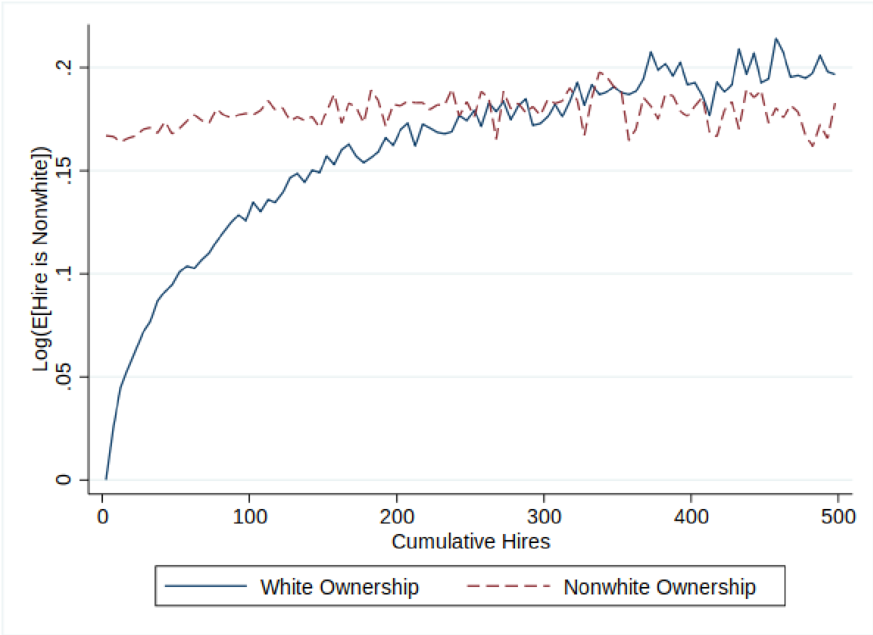
Figure B.2
Share of Workforce in Human Resources Increasing in Employer Size



Note: This figure reports the share of an employer's workforce in human resources-related (HR) occupations. HR occupations include: *administrador* (administrator); *diretor de recursos humanos* (human resources director); *gerente de recursos humanos* (human resources manager); and *gerente de departamento pessoal* (personal department manager). The HR share is calculated for each plant by year combination, then averaged across plant by year combinations weighting by number of hires.

Figure B.3

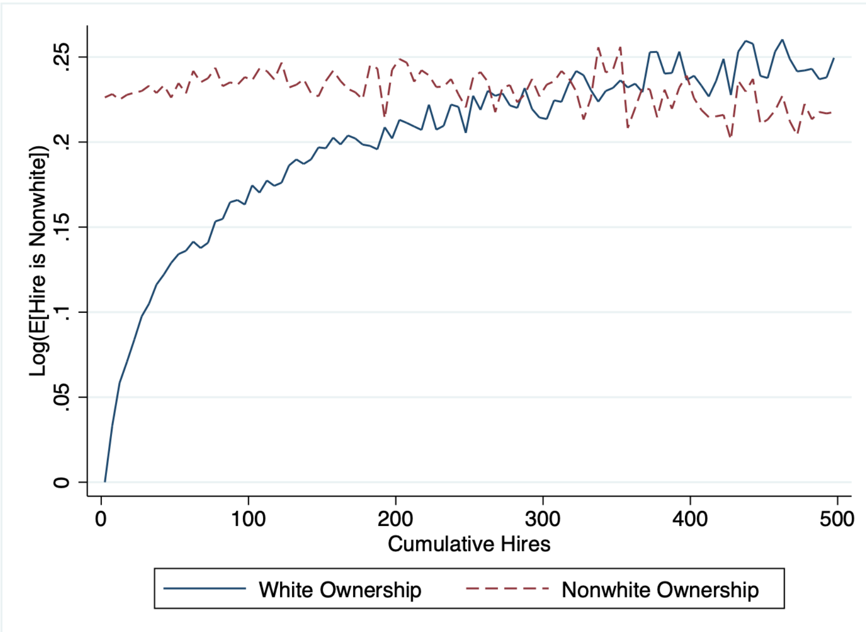
Nonwhite Share of Hires Converges with Cumulative Hires, Controlling for Firm Size



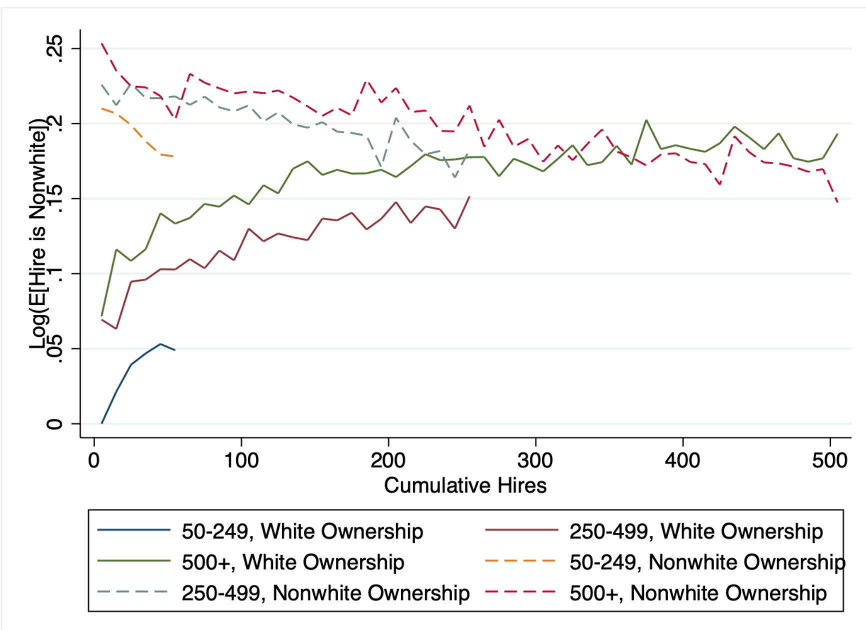
Note: This figure plots the relationship between the racial composition of a firm's hires and its cumulative hires to date. The figure plots the $\eta^{n,r}$ coefficient estimates from equation (3) augmented with HQ establishment size fixed effects, summarizing the relationship between a firm's racial composition of hires, its cumulative hires to date (n), and the race of its founder (r). The establishment size categories are: 1–4, 5–9, 10–19, 20–49, 50–99, 100–249, 250–499, 500-999, and 1000 or more employees. The model is estimated via Poisson quasi maximum likelihood (PQML). The omitted category is the first five hires after the year of entry for firms with white founders. Founder race is inferred from the race of the top-paid manager or employee at entry.

Figure B.4
Nonwhite Share of Hires Converges with Cumulative Hires, by Ownership

(a) Pooled

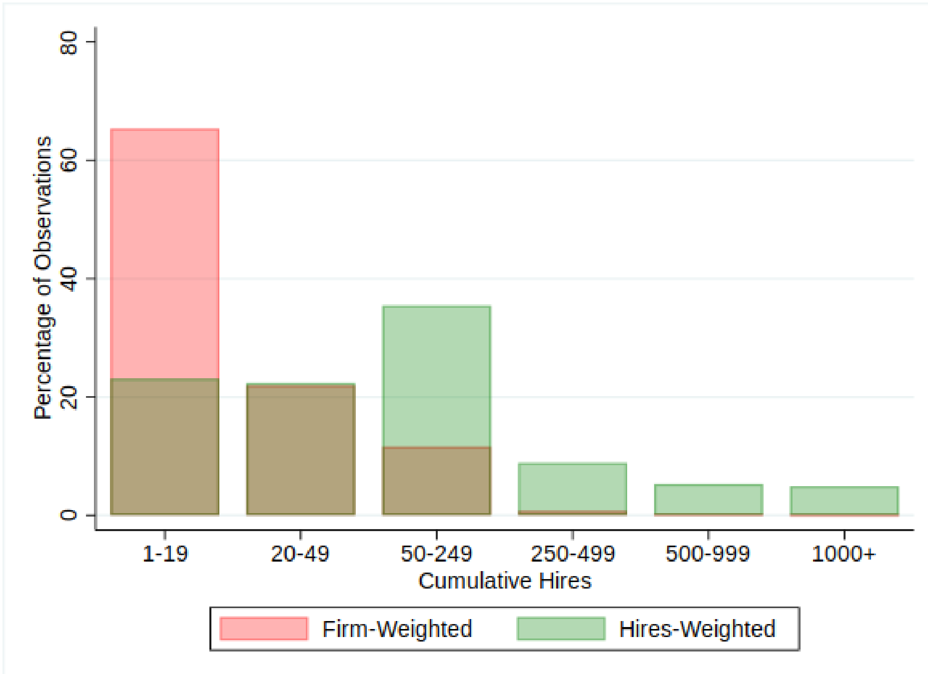


(b) By Total Hires, Balanced



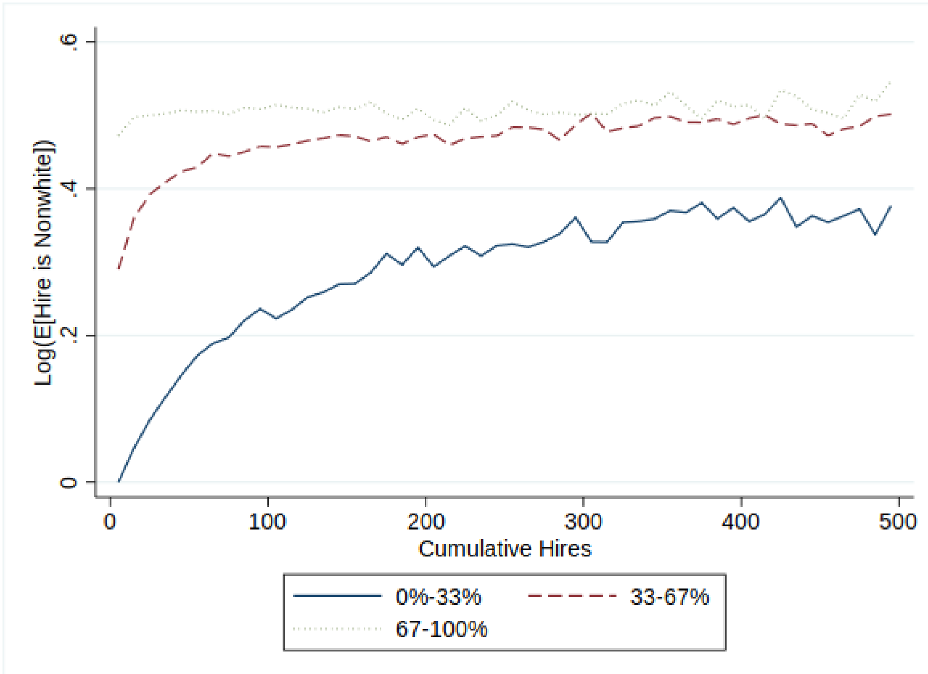
Note: This figure plots the relationship between the racial composition of a firm's hires and its cumulative hires to date. Panel A plots the $\eta^{n,r}$ coefficient estimates from equation (3), summarizing the relationship between a firm's racial composition of hires, its cumulative hires to date (n), and the race of its founder (r). Panel B plots the $\eta^{s,n,r}$ coefficient estimates from equation (4), which allows the relationship between a firm's racial composition of hires, its cumulative hires to date, and the race of the founder to vary with the firm's total observed hires (s). Both models are estimated via Poisson quasi maximum likelihood (PQML). In Panel A the omitted category is the first five hires after the year of entry for firms with white founders. In Panel B the omitted category is the first ten hires after the year of entry for firms with white founders and 50-249 total observed hires. Founder race is inferred from the racial composition of the firm's ownership.

Figure B.5
Cumulative Hires Distribution 5 Years Post-Entry



Note: This figure plots a histogram for the number of cumulative hires by firms 5 years after entry. The sample is limited to firms that remain in the RAIS data 5 years after entry. In the 'Firm-Weighted' bars, each firm is weighted equally. In the 'Hires-Weighted' bars, each firm is weighted by their number of cumulative hires.

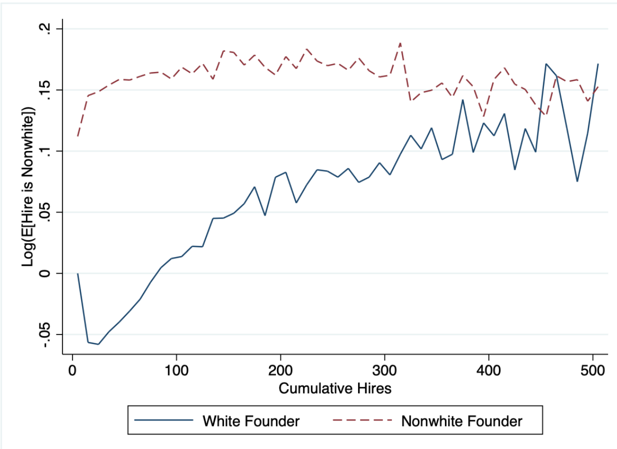
Figure B.6
Nonwhite Share of Hires Converges with Cumulative Hires, New Establishments of Existing Firms



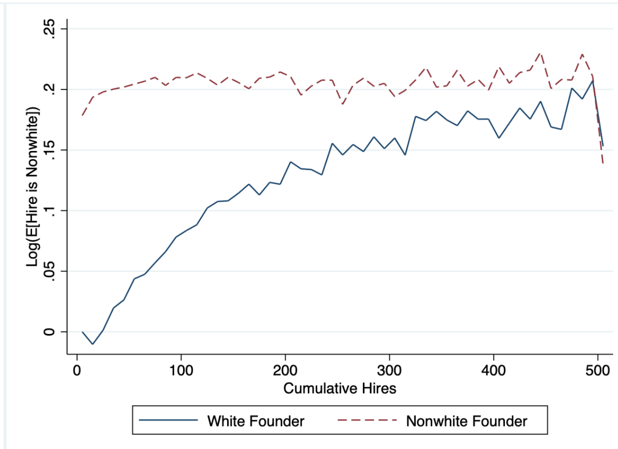
Note: This figure plots the relationship between the racial composition of an entrant establishment's hires and its cumulative hires to date. The figure plots the $\eta^{n,r}$ coefficient estimates from equation (3), summarizing the relationship between an establishment's racial composition of hires, its cumulative hires to date (n), and the racial of employees at incumbent establishments in the firm (r). The model is estimated via Poisson quasi maximum likelihood (PQML). We characterize establishments by the nonwhite share of the firm's incumbent employees, and divide establishments into three categories on this basis: 0-33%, 33-67%, and 67-100%. The omitted category is the first five hires after the year of entry for establishments where the nonwhite share of the firm's incumbent employees is 0-33%.

FIGURE B.7
NONWHITE SHARE OF HIRES BY CUMULATIVE HIRES FOR VARYING FIRM PAY PREMIUMS

(a) Bottom Quintile



(b) Middle Quintile



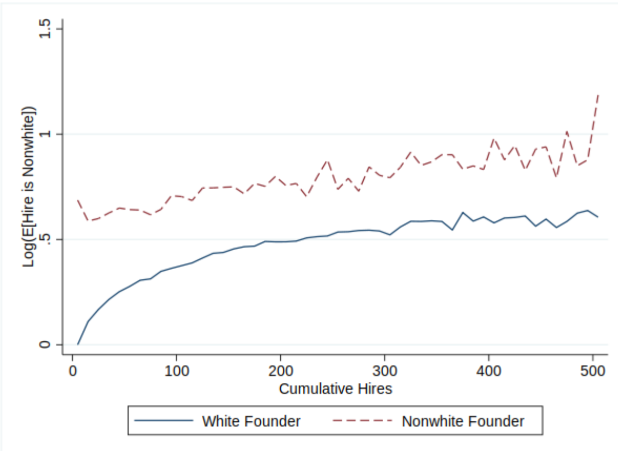
(c) Top Quintile

Note: This figure plots the $\eta^{n,r}$ coefficient estimates from equation (3), summarizing the relationship between an establishment's racial composition of hires, its cumulative hires to date (n) and the race of its founder (r). Firms are grouped by the quintile of their AKM firm effect for white workers. The model is estimated via Poisson quasi maximum likelihood (PQML). In each panel the omitted category is the first hire of establishments with white founders.

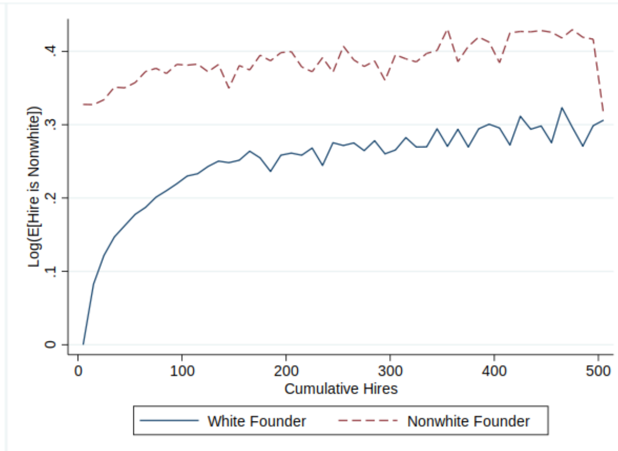
Figure B.8

Nonwhite Share of Hires by Cumulative Hires and Microregion Nonwhite Share

(a) Bottom Quintile



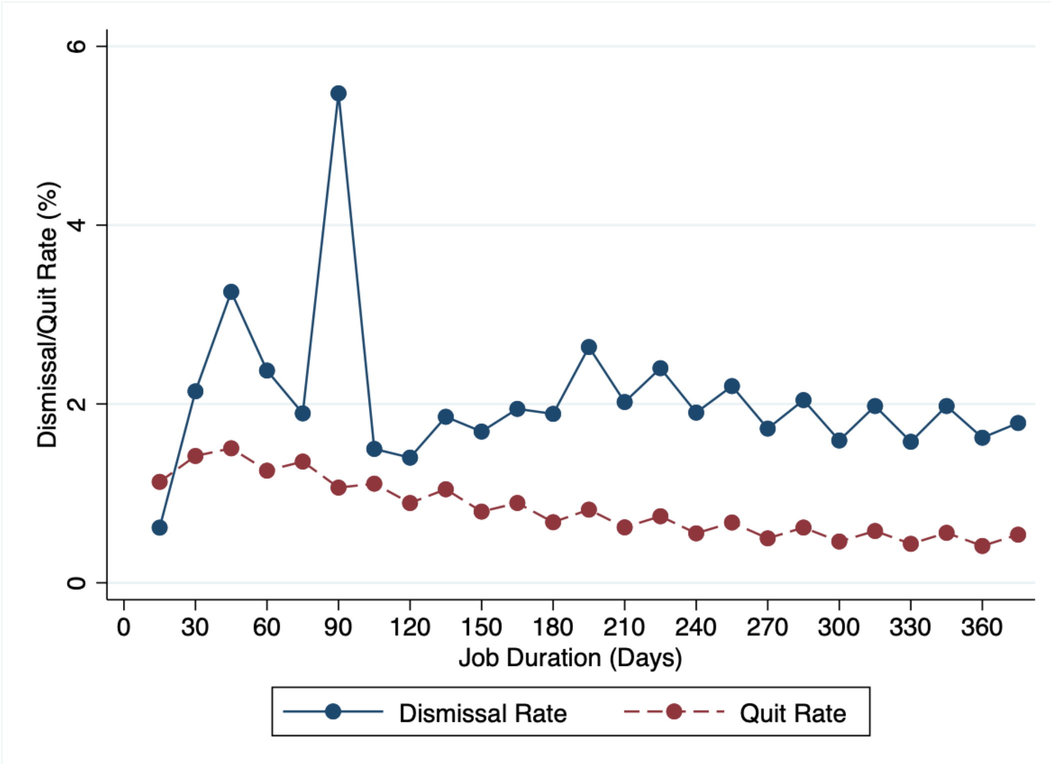
(b) Middle Quintile



(c) Top Quintile

Note: This figure plots the $\eta^{n,r}$ coefficient estimates from equation (3), summarizing the relationship between an establishment's racial composition of hires, its cumulative hires to date (n) and the race of its founder (r). Firms are grouped by quintile for the nonwhite share of hires in the microregion where the firm is located. The model is estimated via Poisson quasi maximum likelihood (PQML). In each panel the omitted category is the first hire of establishments with white founders.

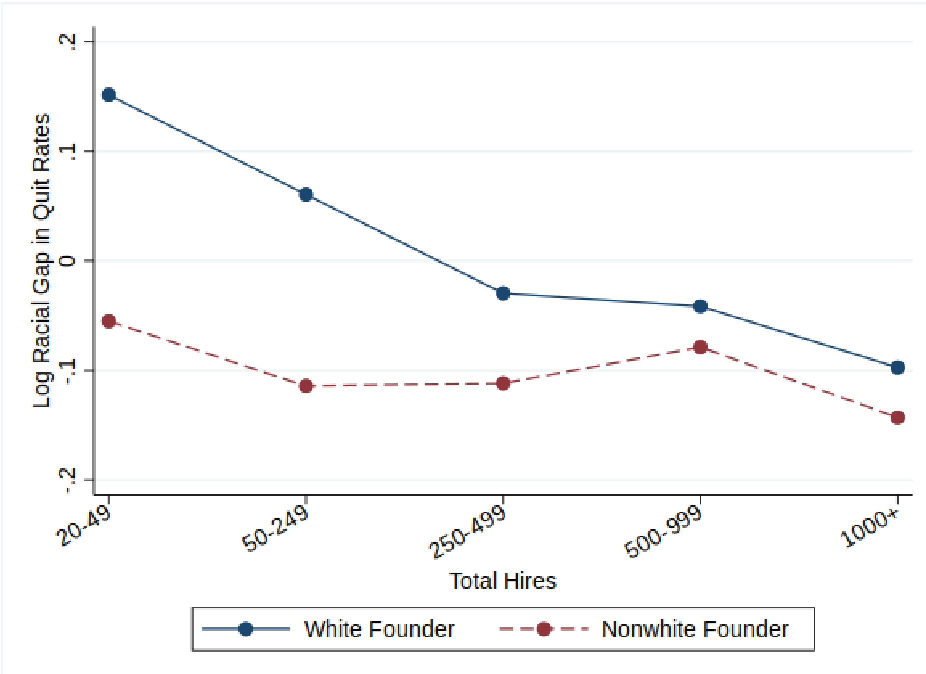
Figure B.9
Dismissal and Quit Rates by Job Tenure



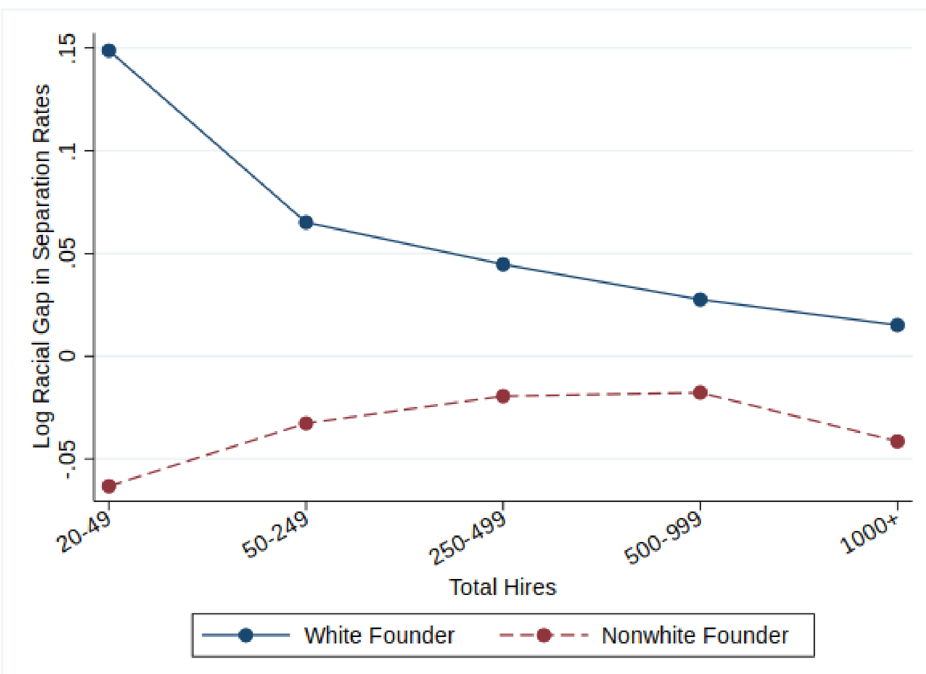
Note: This figure presents dismissal (or employer-initiated or 'involuntary' separations) and quit (or employee-initiated separation) hazard rates as a function of job spell tenure. Tenure is grouped in 15 day periods.

Figure B.10
Racial Disparity in Quit and Separation Rates Rates by Total Hires

(a) Quits



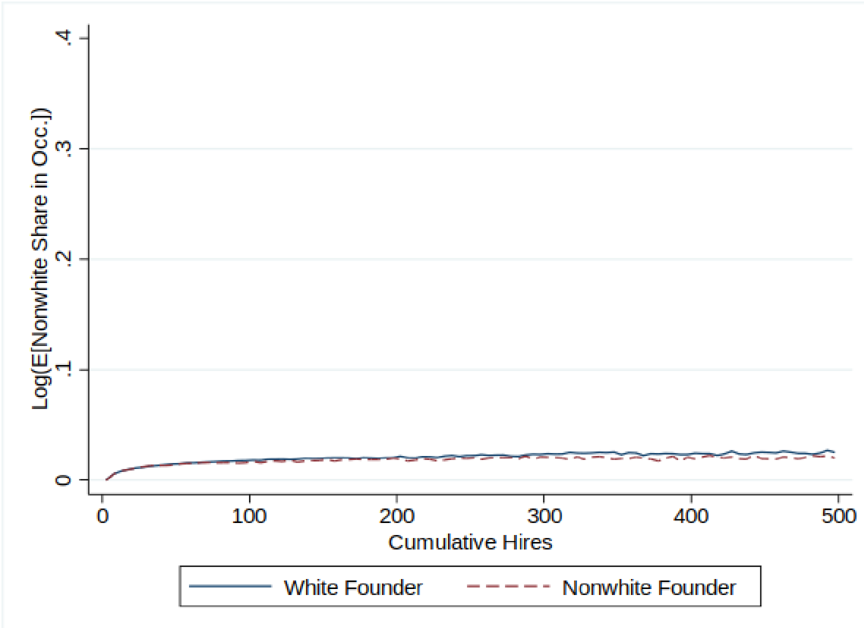
(b) Separations



Note: These figure plot the adjusted, establishment-level nonwhite-white difference in log quit rates (Panel A) and separation rates (Panel B) as a function of founder race and the establishment's total number of observed hires after the year of entry. Establishment-specific racial differences in quit and separation rates are constructed as described in equation (7). The model is estimated via Poisson quasi maximum likelihood (PQML). We limit to establishments with 20 or more hires.

Figure B.11

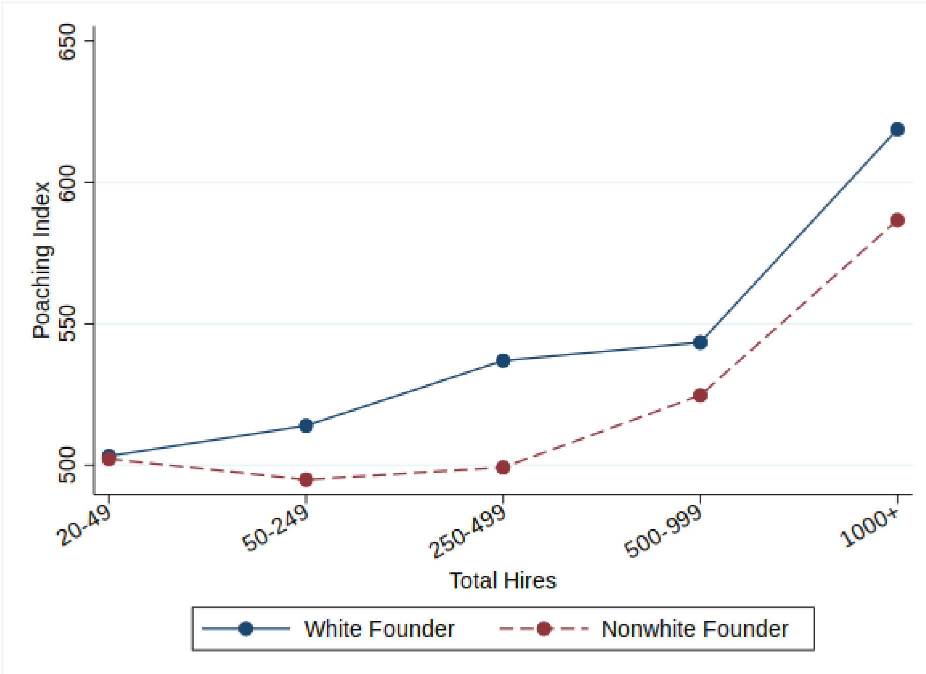
Occupation-Based Predicted Nonwhite Share by Cumulative Hires



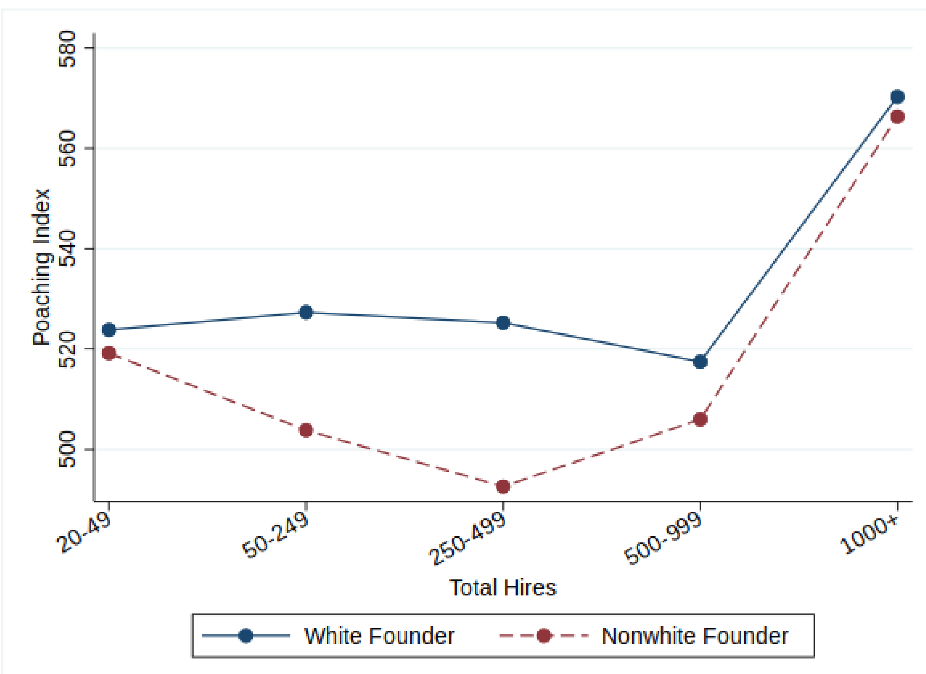
Note: This figure plots the relationship between the occupational mix of a firm's hires and its cumulative hires to date. The figure plots the $\eta^{n,r}$ coefficient estimates from equation (8), summarizing the relationship between a firm's occupational mix of hires, its cumulative hires to date (n), and the race of its founder (r). Occupational mix is characterized by $\bar{\omega}_o$, the nonwhite share of workers hired into that six-digit occupation. The omitted category is the first five hires after the year of entry for firms with white founders. Founder race is inferred from the race of the top-paid manager or employee at entry.

Figure B.12
Poaching Index by Founder Race and Total Hires

(a) White Hires



(b) Nonwhite Hires



Note: This figure reports race-specific average poaching ranks by an establishment's total hires and founder's race. The same of entrant establishments is described in Section 4.2. The poaching rank is defined in Section B.1.

Table B.1
Descriptive Statistics for Displaced Worker Sample

	Dyads (1)	Displaced (2)	Incumbents (3)
Any Link	7.3%		
Linked	3.7%		
Hired	0.074%		
White	26.5%	29.6%	47.0%
Male	41.8%	64.3%	69.3%
Age	34.3	33.5	35.0
Dest. Size			
1–99	63.6%	58.0%	58.3%
100–499	20.0%	24.0%	22.1%
500+	16.5%	18.0%	19.5%
Num. Obs.	11,323,615	39,701	651,306

Source: RAIS, 2013–2017. The "Dyads" column includes pairs of displaced workers matched to potential destinations. The "Incumbents" column describes the population of incumbent workers who are linked to some hired worker via a past coworking relationship.

Table B.2
Referral Effects by Job Seeker and Incumbent Race:
Displaced Worker Sample

	Overall			Race Match
	(1)	(2)	(3)	(4)
True Link	0.167	0.278	0.116	
	(0.012)	(0.020)	(0.014)	
Any Link		0.086	0.052	
		(0.013)	(0.007)	
Race Match × True Link				
Nonwhite / Nonwhite				0.181
				(0.028)
Nonwhite / White				0.016
				(0.018)
White / Nonwhite				0.017
				(0.025)
White / White				0.130
				(0.018)
Dep. Var. Mean.	0.074	0.074	0.074	0.074
Estab. Pair FE	✓		✓	✓
Placebo Link Control		✓	✓	✓
Number of Obs.	11,323,615			

Columns 1–3 presents estimated referral effects under different identifying assumptions. Column 4 reports heterogeneity in referral effects based on the match between the race of the job changer and the race of the linked incumbent. All specifications include controls for worker demographic and human capital characteristics. Column 2 controls for origin and destination establishment effects. Column 4 includes controls for each race match interacted with "Any Link", which indicates observations for which the job changer has either a true coworker or a placebo coworker connection to an incumbent worker at the destination. When reporting the race match, we put the race of the job seeker first and the linked incumbent second. So "White / Nonwhite" indicates a white job seeker is linked to a nonwhite incumbent at the destination. The sample is restricted to worker-target plant dyads for workers that separated from their job during a mass displacement event.

Table B.3
Characteristics of Entrant HQ Establishments

	By Top-Paid Manager			By Ownership		
	Pooled Founders (1)	White Founders (2)	Nonwhite Founders (3)	Pooled Founders (4)	White Founders (5)	Nonwhite Founders (6)
Nonwhite Founder (%)	33.0	0.0	100.0	16.6	0.0	100.0
Persistence						
After 3 Years	65.5	66.4	63.8	63.1	63.6	60.3
After 5 Years	42.3	43.3	40.2	39.4	40.1	35.9
Total Hires						
After 3 Years						
1-19	79.2	79.4	79.0	72.8	73.2	70.6
20-49	14.3	14.3	14.4	17.9	17.8	18.3
50-249	6.1	6.0	6.2	8.6	8.4	10.1
250-499	0.3	0.3	0.4	0.5	0.5	0.8
500-999	0.1	0.1	0.1	0.2	0.1	0.2
1000+	0.0	0.0	0.0	0.0	0.0	0.0
After 5 Years						
1-19	65.4	65.6	64.8	56.8	57.3	54.0
20-49	22.0	21.8	22.4	25.5	25.4	26.3
50-249	11.6	11.6	11.6	15.8	15.5	17.1
250-499	0.8	0.8	0.9	1.3	1.2	1.8
500-999	0.2	0.2	0.3	0.4	0.4	0.6
1000+	0.1	0.1	0.1	0.2	0.2	0.2
Number of Firms	2.27m	1.52m	0.75m	591k	493k	98k

This table reports summary statistics for entrant HQ establishments in the *Relação Anual de Informações Sociais* (RAIS) data for the years 2003–2017. In columns 1 through 3 we infer the race of the firm's founder using the race of the top-paid manager (or top-paid employee if there is no manager present) in the year of entry. In columns 4 through 6 we infer the race of a firm's founder using the racial composition of ownership. We classify firms where more than 50% of ownership is white as having a white founder and firms where more than 50% of ownership is nonwhite as having a nonwhite founder.