The Dynamic Effects of Co-Racial Hiring

Conrad Miller*

Ian Schmutte†

January 2023

Abstract

In Brazil, firms' later hires are more likely to be nonwhite than early hires for the same job. We argue that this pattern reflects racial disparities in entrepreneurship and co-racial hiring: firms are more likely to hire from groups already well-represented at the firm, though with some decay. At entry, firms with white founders are about 30% less likely to hire nonwhite employees than comparable firms with nonwhite founders. After 400 hires, these firms nearly converge in their composition of subsequent hires. Yet few firms reach this scale. Within-firm racial differences in dismissal rates follow an analogous pattern. We provide suggestive evidence that referral hiring can at least in part account for our findings.

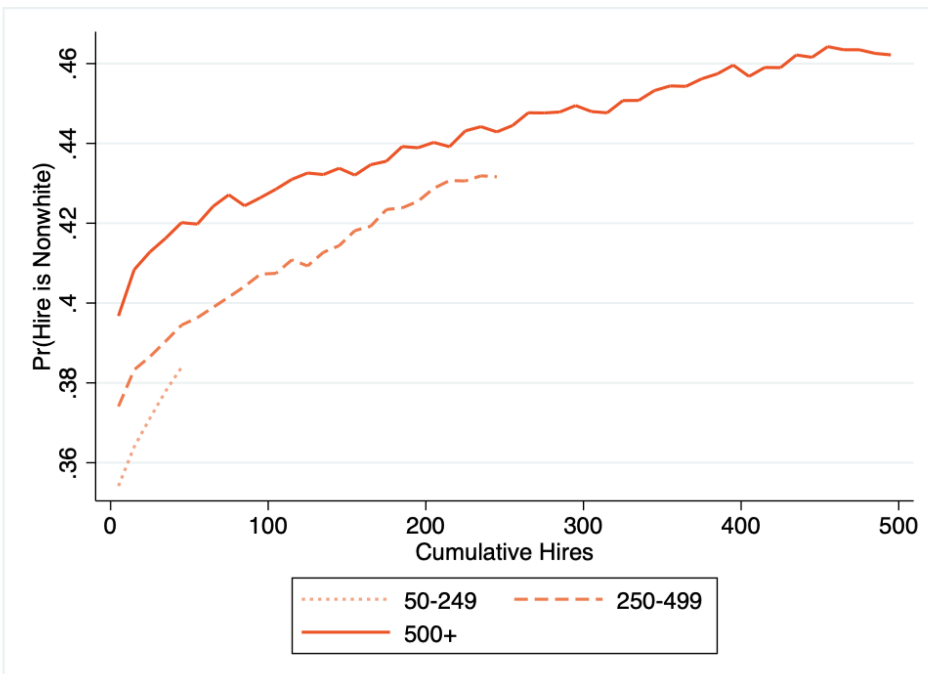
*Haas School of Business, University of California, Berkeley and NBER (email: ccmiller@berkeley.edu)

†Terry College of Business, University of Georgia (email: schmutte@uga.edu). We thank Joaquín Fuenzalida for excellent research assistance. We thank Kathryn Anderson, Iwan Barankay, Alex Bartik, Sydnee Caldwell, Jonathan Leonard, Louis-Pierre Lepage, David Levine, Ricardo Perez-Truglia, Isaac Sorkin, Guo Xu, and numerous seminar participants for comments. This paper is a significant revision of and supersedes the previously circulated manuscript "Racial Inequality and the Dynamic Role of Referral Hiring: Evidence from Brazil". Miller gratefully acknowledges financial support from the Upjohn Institute and the UC Berkeley Institute for Research on Labor and Employment. Schmutte gratefully acknowledges financial support from the Terry College Institute for Leadership Advancement and the Bonbright Center for the Study of Regulation.

1 Introduction

In Brazil, whites earn more than nonwhites and are less likely to be unemployed (IBGE, 2021). We document and examine a related stylized fact: as a firm's cumulative hires increase, so does the probability that its next hire is nonwhite. For example, among formal sector firms that make at least 500 hires during our study period, the nonwhite share of first hires is 40%, increases to 46% by the 400th hire, and stagnates thereafter (see Figure 1). This change is substantial: a five percentage point increase in the nonwhite share of the formal sector workforce would eliminate racial differences in formal employment rates.[1] We show that this stylized fact is not explained by changes in job characteristics, including detailed occupation and firm size. There are substantive changes in labor demand (or supply) by worker race over a firm's life cycle.

Figure 1
Nonwhite Share of Hires Increases with Cumulative Hires



Note: These figures plot the relationship between the racial composition of a firm's hires and its cumulative hires to date. The sample includes entrant firms that make 50-249, 250-499, or 500 or more hires during our sample window, 2003–2017.

Source: Relação Anual de Informações Sociais (RAIS) data, 2003–2017

We argue that the positive relationship between firms' cumulative hires and nonwhite share of hires is driven by two factors: racial disparities in entrepreneurship and co-racial hiring. By co-racial hiring we mean mechanisms or hiring practices that favor candidates from whichever racial group is already well-represented among incumbent employees at the firm. Possible mechanisms

[1] This statistic refers to 18–65 year olds with at least one year of potential work experience and is based on Pesquina Nacional for Amostra de Domicilios (PNAD) household survey data from 2003 to 2015. See Section 2.1 for more details.

include referral hiring or a tendency for managers to hire employees from their own racial group, possibly due to prejudice, production complementarities, or screening discrimination. Regardless of the specific mechanism, co-racial hiring generates persistence in the racial composition of a firm's hires; new firms with disproportionately white (nonwhite) incumbent employees will tend to hire white (nonwhite) workers. Figure 1 suggests that this persistence *decays* over time. As a firm's cumulative hires increase, the racial composition of its hires becomes less dependent on the initial composition of its employees and more closely tied to the composition of the local labor market. Given that whites are twice as likely to start new firms as nonwhites (IBGE, 2021), the average firm becomes more nonwhite in composition over its life cycle.

Consistent with co-racial hiring, we find that the racial composition of hires for firms with white and nonwhite founders differs substantially at entry. For the same local labor market and occupation, early hires at firms with white founders are about 30% less likely to be nonwhite than early hires at firms with nonwhite founders. Consistent with decay, the composition of hires for these two sets of firms nearly converge as their cumulative hires increase. At firms with a white founder the nonwhite share of hires increases sharply with cumulative hires; at firms with nonwhite founders the nonwhite share of hires *decreases* with cumulative hires. After 400 hires, the next hire at firms with white founders are only 4% less likely to be nonwhite than the next hire at firms with nonwhite founders. We find a similar pattern for new establishments in existing firms where we categorize the establishment's initial social conditions by the racial composition of the firm's incumbent employees in preexisting establishments. More generally, the dispersion across firms in their racial composition of hires decreases as cumulative hires increase.

Dismissals follow an analogous pattern. 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. Racial differences in dismissal rates diminish over the firm's life cycle. We also find that the racial composition of a firm's hires predicts racial differences in dismissal rates: firms that hire disproportionately white employees are more likely to dismiss their nonwhite hires. These findings are consistent with several forms of co-racial hiring. Firms may make fewer mistakes in screening co-racial candidates and co-racial hires may be less likely to be marginal hires (Benson et al., 2022).

Though the importance of founder race diminishes with cumulative hires, few firms reach the scale at which the nonwhite share of hires stagnates. We estimate that, for the average entrant firm in our sample, the nonwhite share of hires is 4% below the steady state value.

We provide evidence that referral hiring is a form of co-racial hiring that can explain our findings, at least in part. We show that referral networks are racially segregated and socially-connected hires are less likely to be dismissed. We do not observe referrals directly. Instead, we proxy for social connections by identifying pairs of workers who worked together in the past.[2] Following Eliason

[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.

et al. (2020), we measure the effect of social connections on where workers sort by comparing the destinations of workers who 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 who 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. We also find that recent hires with preexisting social connections at their employers are dismissed at lower rates than recent hires with placebo connections.[3]

We consider alternative explanations for our findings that do not involve co-racial hiring. One hypothesis is that our findings reflect the evolution of labor *supply* rather than labor demand; in particular, jobseekers may prefer to work with co-workers of the same race. However, worker mobility patterns suggest that white and nonwhite workers have similar preferences over employers, at least as a function of founder race and cumulative hires (Sorkin, 2018; Bagger and Lentz, 2018). A second hypothesis is that mature firms are more likely to have a human resources (HR) department, which makes hiring and turnover more equitable between groups by formalizing hiring and evaluation processes (Dobbin, 2009). Yet our findings are similar for both firms that do and do not employ workers in HR occupations.

A growing body of literature documents referral hiring (Fernandez et al., 2000; Fernandez and Sosa, 2005; Petersen et al., 2000; Dustmann et al., 2016) and bias among hiring managers (Giuliano et al., 2011; Åslund et al., 2014; Benson et al., 2022) as important sources of co-racial hiring. Relative to prior work, we emphasize dynamic, which are critical for understanding the implications of co-racial hiring in the aggregate or in the long-run for a given firm. Co-racial hiring can, in principle, lead to tipping or extreme segregation (Becker, 1957; Schelling, 1971; Lang, 1986; Pan, 2015). Co-racial hiring may also decay, leading to at least partial convergence. For example, the co-racial hiring that referral hiring produces will decay if referral networks are not perfectly segregated (Rubineau and Fernandez, 2013, 2015) or employees who themselves were not referrals eventually generate referrals of their own. Our findings are consistent with decay. However, we find that the average entrant firm does not make enough hires to reach convergence, suggesting that co-racial hiring dampens aggregate relative demand for nonwhite workers.

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 within the same local labor market. In Brazil, Dias and Rocha (2021) document a similar tendency for business owners to hire workers of the same race and find that racial wage disparities are smaller in firms with nonwhite ownership. Kerr and Kerr (2021) document co-ethnic hiring by immigrant entrepreneurs in the United States. We emphasize that the role of founder race dissipates over the firm life cycle, though the convergence

[3]This finding mirrors prior research that finds that referral hires have lower turnover rates (Brown et al., 2016; Topa, 2019).

process is slow.

Finally, 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 papers that use similar approaches. Hensvik and Skans (2016) compare true and placebo coworkers to infer the characteristics of workers who receive and provide referrals, and the effects of being referred on job outcomes. Caldwell and Harmon (2019) study the effect of coworker networks on job mobility and earnings and compare outcomes based on 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. We identify a sharp discontinuity in hiring for workers who actually overlap to those who almost 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.

Section 2 describes the Brazilian context and employer-employee data that form the basis of our study. In Section 3 we show that the positive relationship between firms' cumulative hires and nonwhite share of hires cannot be explained by observable job characteristics. Section 4 describes our model of co-racial hiring and its predictions. Section 5 presents supporting evidence. In Section 6 we show that referral hiring is a relevant form of co-racial hiring in our context. In Section 7 we address alternative explanations of our findings that do not involve co-racial hiring. Section 8 concludes.

2 Context and Data

Like the United States, Brazil's labor market exhibits significant racial disparities in wages and unemployment and workplace segregation (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 basis of race (Machado et al., 2019). 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 includes detailed information on both workers and their employment contracts.

2.1 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 develop 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.[4]

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. 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.

To provide more context, we summarize data from the Pesquina Nacional for Amostra de Domicilios (PNAD) between 2003 and 2015.[5] 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.

About 48% of working-age Brazilian adults 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. Unemployment rates are 25% to 30% higher for nonwhites.

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. 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, whites have more years and education and receive wages that are 20 to 30 log points higher than wages received by nonwhites. About 80% of private sector employees report having a valid *carteira do 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.

2.2 RAIS Employer-Employee Data

Our analysis uses RAIS data from 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 education, race, and gender as

[4]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).

[5]Our summary of PNAD statistics mirrors Gerard et al. (forthcoming).

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

Note: 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.

reported by the employer. We identify the race for each individual using their modal reported race across all contract-years for which they appear in the data.[6] 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 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.

We limit the sample to worker-firm-year observations for men and women on private sector, indeterminate-length contracts. We also limit most of our analysis to jobs with entrant firms, which we define as firms that hire their first employee observed in the RAIS data during our sample window. 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 are left with a sample of about 3.2 million HQ establishments.[7] In some of the analysis we look at new establishments that are subsidiaries of preexisting firms. We identify about 700 thousand new establishments from preexisting firms.

For entrant firms, 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.[8] Second, when possible, we infer the race of a firm's founder using the racial composition of ownership (see Section 2.3 below for a description of the ownership data). We classify firms as having a white founder when we can identify more than 50% of ownership as white and as having a nonwhite founder when we can identify more than 50% of ownership as nonwhite.[9]

Table 2 describes our main sample of entrant firms. Notably, using either classification, we find that entrant firms with white and nonwhite founders are similar in terms of their size, industry, and survival rates.

There are significant racial disparities in wages in the RAIS data (see Appendix Table C.1 for descriptive statistics). There is a 20 log point (22%) raw wage gap between white and nonwhite workers. This gap is partially explained by differences in education: white workers are 7.3 percentage points more likely to be college graduates. Recently hired workers are more likely to be nonwhite. The raw racial wage gap among new hires is smaller than the overall gap, at 12 log points (13%). Racial differences between recently hired workers are not meaningfully different when we

[6]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.

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

[8]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.

[9]The first method may inflate the nonwhite share of founders but covers a substantially larger set of firms. In calculating the white and nonwhite share of ownership, we include owners that we cannot match to the RAIS data in the denominator.

Table 2
Characteristics of Entrant HQ Establishments

	By Top-Paid Manager			By Ownership		
	Pooled	White Founders	Nonwhite Founders	Pooled	White Founders	Nonwhite Founders
	(1)	(2)	(3)	(4)	(5)	(6)
Nonwhite Founder (%)	31.8	0.0	100.0	17.2	0.0	100.0
Total Hires						
1-19	59.4	59.7	58.6	50.7	51.2	47.6
20-49	24.6	24.4	25.1	27.6	27.5	28.2
50-249	14.3	14.3	14.5	18.9	18.7	20.4
250-499	1.2	1.2	1.3	1.9	1.8	2.5
500-999	0.4	0.4	0.4	0.7	0.6	1.0
1000+	0.1	0.1	0.2	0.3	0.3	0.4
Survival						
After 3 Years	51.7	53.0	48.9	49.6	50.4	45.7
After 5 Years	33.2	34.5	30.6	30.8	31.5	27.1
Industry (%)						
Manufacturing	9.8	10.5	8.5	10.4	10.9	8.1
Construction	6.4	5.9	7.6	7.8	7.1	11.3
Commerce	45.1	44.8	45.8	39.1	38.7	41.1
Transport, Storage, and Mail	6.2	6.6	5.5	5.8	5.9	5.2
Accommodation and Meals	9.2	8.9	10.0	8.5	8.7	7.6
Professional Activities	3.6	3.7	3.2	4.7	4.8	4.3
Administrative Activities	6.4	6.3	6.6	6.9	6.8	7.2
Health and Social Services	2.3	2.3	2.2	3.4	3.7	2.4
Other	11.0	11.0	10.6	13.4	13.4	12.8
Number of Firms	3.21m	2.19m	1.02m	847k	701k	146k

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. *Total Hires* refers to total hires that we observe during our study period. *Survival* refers to whether a firm persists in the RAIS data.

limit the sample to entrant firms.

A key limitation of the RAIS data is that it excludes informal firms and informal employment contracts. Over our study period, 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). Given data constraints, our conclusions throughout apply only to formal contracts. Our key findings hold across industries, which vary in informality levels.

2.3 CNPJ Ownership Data

In one approach to inferring founder race, 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.

3 Job Characteristics Cannot Explain Life Cycle Pattern

This paper is motivated by a stylized fact: firms' later hires are more likely to be nonwhite than their early hires. One potential explanation is that the types of job vacancies that firms fill changes over the life cycle. For example, firms may tend to first hire employees in managerial or professional occupations—positions disproportionately held by white workers, who have more years of formal education on average—and later hire for other positions with lower education requirements. In this section we examine whether job characteristics, including detailed occupation and contemporaneous firm size, can explain the stylized life cycle pattern.

Let j index firms and let h index hires within a firm by start date where, for firm j, $h \in \{1, ..., H_j\}$. We estimate regression models of the form

$$\log(E(\text{NONWHITE}_{jh}|\cdot)) = \sum_n \eta^n \times \mathbb{1}_{\{N(j,h)=n\}} + \tau_{t(j,h)} + \psi_j + X_{jh} + \epsilon_{jh} \tag{1}$$

via Poisson quasi maximum likelihood (Correia et al., 2020), where each observation is a new hire.[10]

[10]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 use a proportional model because we expect effects to be proportional to the racial composition of the local labor market. We also estimate η^n coefficients using a linear probability model and obtain similar results. See Appendix C for details.

NONWHITE_{jh} is an indicator for whether hire h at firm j is nonwhite. $N(j,h)$ groups firm j's hires into bins.[11] We limit the estimation sample to firms' first 500 hires and group hires into increments of ten: hires 1–10, 11–20, 21–30, and so on, up to 491–500. ψ_j are firm fixed effects, $\tau_{t(j,h)}$ are year fixed effects, and X_{jh} is a vector of additional controls for job characteristics. We vary this set of controls across specifications. We exclude the inferred founder from the new hires we consider and when measuring cumulative hires. The omitted category is the first increment of hires. The η^n coefficients have a proportional interpretation: they measure the proportional increase in the probability that a hire is nonwhite relative to initial hires.

We plot the η^n coefficient estimates for four specifications in Panel A of Figure 2.[12] The baseline specification, depicted in blue, includes firm fixed effects and year fixed effects, but no additional controls. The probability that a hire is nonwhite increases by 7 to 8 log points from the first bin of hires to about the 400th hire, and plateaus thereafter.

The second specification (red) includes 6-digit occupation fixed effects. The inclusion of occupation fixed effects moderately attenuates the η^n coefficient estimates. Conditional on occupation, the probability that a hire is nonwhite increases by 5 to 6 log points over firms' first 400 hires.

Occupation fixed effects alone may miss important variation in job characteristics if jobs are coded differently across firms or if the nature of jobs associated with specific occupations varies across firms. To address this concern, the third specification (green) replaces firm and occupation fixed effects with firm by occupation interaction fixed effects. The coefficients are virtually unchanged.

Even within the same occupation by firm pair, the tasks required for a job may vary over a firm's life cycle. In particular, the nature of what is nominally the same job may differ when a firm is small compared to when the firm is large. The fourth specification (orange) further interacts the firm by occupational fixed effects with fixed effects for the firm's contemporaneous size. We bucket firm size into the following buckets by number of employees: 1–19, 20–49, 50–249, and 250–500. Controlling for firm size does not meaningfully change the coefficient estimates.

Another concern with interpreting the pattern illustrated in Panel A of Figure 2 is that the set of firms that contribute to the estimation of η^n coefficients varies with n. We next estimate equation (1) for a balanced panel of firms. We also allow the η coefficients to vary with a firm's total observed hires. Specifically, we estimate

$$\log(E(\text{NONWHITE}_{jh}|\cdot)) = \sum_s \sum_n \eta^{s,n} \times \mathbb{1}_{\{S(j)=s\}} \times \mathbb{1}_{\{N(j,h)=n\}} + \tau_{t(j,h)} + \mu_{m(j)} + \omega_{o(j,h)} + \epsilon_{jh}, \quad (2)$$

where $\omega_{o(j,h)}$ are occupation fixed effects and $S(j)$ categorizes firms by their total observed hires: 50–249, 250–499, and 500 or more. We include microregion fixed effects, $\mu_{m(j)}$, which we use to approximate local labor markets, rather than firm fixed effects so that we can compare levels across $S(j)$ categories. We restrict estimation to hires 1–50 for firms with 50–249 total observed hires,

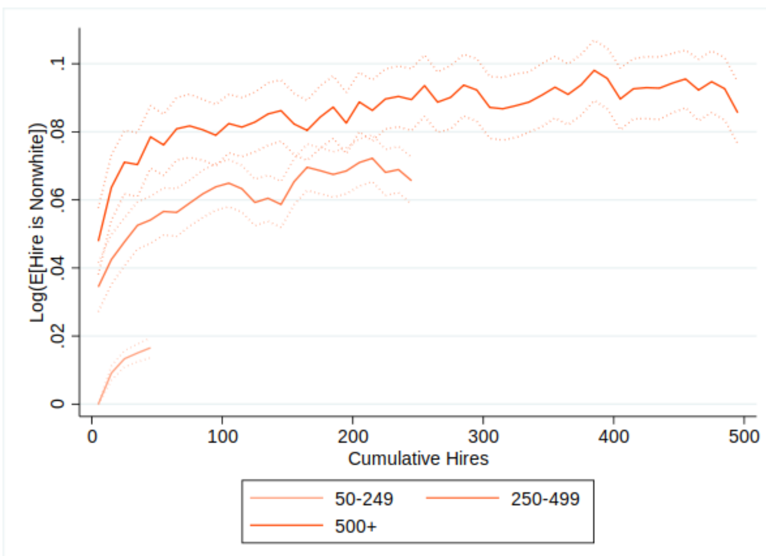
[11] For new hires with the same start date, ties are broken randomly.

[12] See Appendix Table C.2 for statistics describing the estimation sample.

Figure 2
Nonwhite Share of Hires Increases over Life Cycle Within Job

(a) Controlling for Job Characteristics

(b) Balanced Panel



Note: Panel A plots the η^n coefficient estimates from equation (1), summarizing the relationship between a firm's racial composition of hires and its cumulative hires to date (n). The figure includes estimates for four specifications. The baseline specification (blue) includes firm fixed effects and year fixed effects, but no additional controls. The second specification (red) also includes 6-digit occupation fixed effects. The third (green) and fourth (orange) specifications replace firm and occupation fixed effects with firm by occupation fixed effects and firm by occupation by contemporaneous firm size fixed effects, respectively. The figure includes point wise 95% confidence intervals for the baseline specification and the fourth, most saturated specification. Standard errors are clustered at the firm level. Panel B plots the $\eta^{s,n}$ coefficient estimates from equation (2), which allows the relationship between a firm's racial composition of hires and its cumulative hires to date to vary with the firm's total observed hires (s). The figure includes point wise 95% confidence intervals, where standard errors are clustered at the firm level. All models are estimated via Poisson quasi maximum likelihood. We exclude the inferred founder from the new hires we consider and when measuring cumulative hires. In Panel A the omitted category is the first ten hires after the year of entry. In Panel B the omitted category is the first ten hires after the year of entry 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 or more total observed hires. This restriction maintains a balanced sample of firms contributing to the estimation of $\eta^{s,n}$ coefficients.

The results are presented in Panel B of Figure 2.[13] There is a similar increasing relationship for each $S(j)$ firm category. Interestingly, there are large intercept differences between categories. For example, initial hires at firms that we observe making 500 or more hires are 5 log points more likely to be nonwhite than initial hires at firms that we observe making between 50 and 249 hires. We discuss this pattern in more detail in Section 5.1.[14]

In summary, we find that observable job characteristics, including detailed occupation and firm size, can only explain a small share of the positive relationship between a firm's cumulative hires to date and its nonwhite share of hires.

We use the η^n coefficients depicted in Panel A of Figure 2 to calculate the expected deviation between a firm's realized nonwhite share of hires and its plateau nonwhite share (which we interpret below as a steady state), averaged across entrant firms. In particular, we use the η^n coefficients from the most saturated model to calculate the difference between the probability that each hire is nonwhite and the probability at 400 hires, average across all hires for a given firm, and then average across firms. The end result is a weighted average of η^n coefficients where the weights depend on the distribution of total hires across firms. More concretely, consider firm j that makes n_j hires. The expected deviation for firm j is

$$\Delta_j = \bar{\eta} - \frac{1}{n^j} \sum_{i}^{n_j} \eta^i,$$

where $\bar{\eta}$ is the steady state value.

We calculate that, for the average entrant firm, the nonwhite share of their hires is 4% below the plateau value.

3.1 Heterogeneity in Life Cycle Pattern

We explore heterogeneity by firm pay premiums (Abowd et al., 1999), by the racial composition of the local labor market, and by industry in the life cycle pattern depicted in Figure 2. We find that the increasing relationship between nonwhite share and cumulative hires is more pronounced for high-paying firms and in microregions where the nonwhite share of the population is relatively small.

3.1.1 By Firm Pay Premiums

Gerard et al. (forthcoming) find that nonwhite workers are underrepresented at high-paying firms in Brazil. This could in part reflect that high-paying firms begin further from their steady state

[13] See Appendix Table C.3 for statistics describing the estimation sample.

[14] Holzer (1998) and Miller (2017) document that Black workers sort to larger employers in the United States.

nonwhite share of hires. We measure firm pay premiums using 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. We estimate the AKM model separately by region, restricting the sample to dominant job contract-years for workers between 20 and 60 years of age. We control for time-varying worker 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). See Appendix A.1.4 and Abowd et al. (2002) for additional details regarding the estimation and identification methods.

Within each microregion, we divide firms into quartiles by their estimated firm pay premium. We then estimate (1) separately by quartile. Here we limit estimation to the first 100 hires because low-paying firms are particularly unlikely to make many hires.

Panel A of Figure 3 plots the η coefficients separately by quartile. The slope is increasing in firm pay premium. By the 100th hire, the nonwhite share of hires has increased by about 2.5% for the bottom quartile by pay premium; for the top quartile, the nonwhite share of hires as increased by 4.5%.

3.1.2 By Microregion Racial Composition

The racial composition of workers varies dramatically across local labor markets. We divide microregions into quartiles by the nonwhite share of hires in the microregion. In the four quartiles, the nonwhite shares of hires are 10.2%, 28.2%, 47.6%, and 74.6%.

We estimate (1) separately by microregion quartile. Panel B of Figure 3 plots the η coefficients for each quartile. The increasing relationship is more prominent in local labor markets where the nonwhite share of the population is relatively small. In the bottom quartile of microregions, where the nonwhite share of workers is lowest, the probability that a new hire is nonwhite increases by over 20% from the first hire to the 400th hire. By contrast, in the top quartile the relationship is essentially flat.

3.1.3 By Industry

We divide firms into six industries: manufacturing, construction, commerce, transport, storage and mail, accommodation and meals, and services. We estimate (1) separately by industry. Again we limit estimation to the first 100 hires.

Appendix Figure C.1 plots the η coefficients. The slope is similar across industry categories.

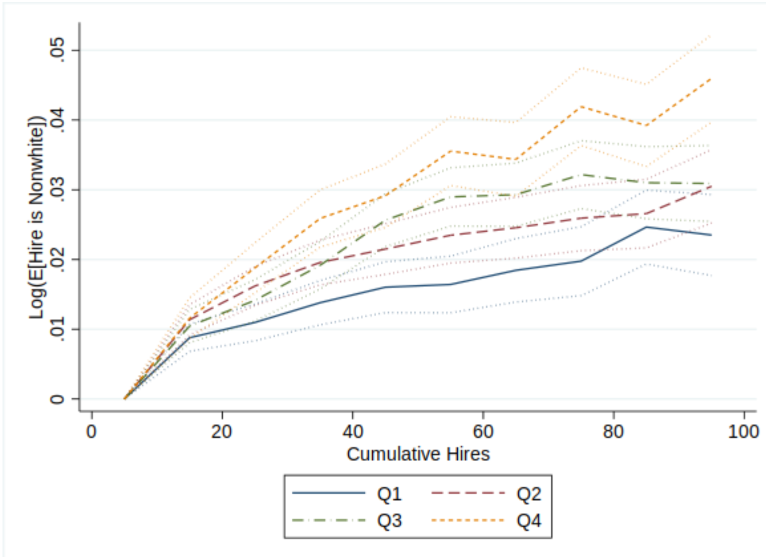
4 A Model of Co-Racial Hiring

For the same job, firms are more likely to hire nonwhite workers later in the firm life cycle. We propose that this stylized fact is driven by two factors: racial disparities in entrepreneurship and co-racial hiring. Firms tend to hire employees from racial groups that are already well-represented

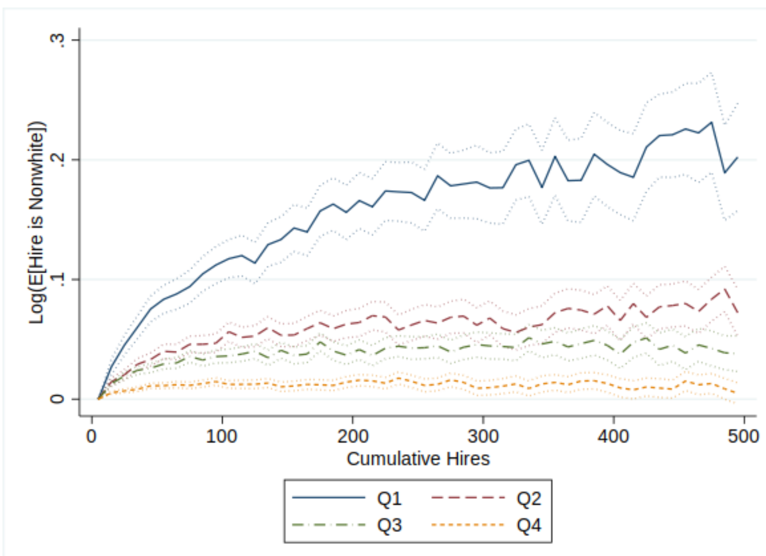
Figure 3

Heterogeneity in Life Cycle Pattern by Firm Pay Premium and Microregion Composition

(a) By Firm Pay Preimum



(b) By Microregion Nonwhite Share



Note: Panel A plots the η^n coefficient estimates from equation (1), summarizing the relationship between an establishment's racial composition of hires and its cumulative hires to date (n). Firms are grouped by the quartile of their AKM firm effect for white workers, stratified by microregion. The model is estimated via Poisson quasi maximum likelihood. We exclude the inferred founder from the new hires we consider and when measuring cumulative hires. The omitted category is the first bin of hires in each quartile of firms. Panel B plots the relationship between the racial composition of a firm's hires and its cumulative hires to date. The figure plots the η^n coefficient estimates from equation (1), summarizing the relationship between a firm's racial composition of hires and its cumulative hires to date (n), estimating separately for four groups of microregions. Microregions are grouped into quartiles by the nonwhite share of hires in the microregion. The model is estimated via Poisson quasi maximum likelihood. The omitted category is the first bin of hires in each quartile group.

at the firm. If this tendency is not too severe, this dependence on initial conditions decays over time, and the racial composition of a firm's hires in steady state is determined by external market conditions. Since founders are disproportionately white, for the average firm the nonwhite share of hires increases over the firm's life cycle.

There are several reasons that the racial composition of a firm's hires may depend on the composition of its incumbent employees. Given homophily in referral networks, incumbent employees are more likely to refer same-race candidates, and firms may have more information about the match quality of referral candidates. Hiring managers may be better able to screen same-race candidates. Managers may have discriminatory tastes and prefer to work with same-race co-workers.[15] Same-race co-workers may be more productive due to production complementarities.

We describe a simple model of co-racial hiring in Appendix B, which we summarize here. The model draws heavily from Benson et al. (2022). Each hire is associated with a randomly selected incumbent employee. One can interpret the selected incumbent as the employee providing the referral or the hiring manager making the decision. Hence, the incumbent's characteristics—their tastes, ability to screen, production function—influence the race of the next hire. The incumbent observes candidate productivity with noise and makes a hiring decision based on that noisy signal. The probability that hire h and firm j' is nonwhite follows

$$Pr(\text{NONWHITE}_{jh}) = f(\theta,\beta) \times \theta + g(\theta,\beta) \times \pi_{jh} \tag{3}$$

where θ reflects the nonwhite share of candidates in the external market and π_{jh} is the nonwhite share of incumbent employees at the firm at the time of hire h. The parameter β measures the degree of co-racial hiring, where $1+\beta$ is the proportional increase in the probability that co-racial candidates are deemed qualified relative to out-group candidates. For simplicity, we assume β is the same for both groups. The functions $f(\theta,\beta)$ and $g(\theta,\beta)$ take values between zero and one and satisfy $f(\theta,0) = 1$ and $g(\theta,0) = 0$.

All forms of co-racial hiring we have discussed may produce racial differences in dismissal rates, where firms are less likely to dismiss co-racial hires (Topa, 2019; Benson et al., 2022). If co-racial hiring is driven by an informational advantage where firms have more information about the match quality of co-racial candidates ex-ante, co-racial hires may be less likely to be a poor fit ex-post. With taste-based discrimination or production complementarities, co-racial hires may be less likely to be marginal. In the model, some hires are dismissed during a probationary period because they are less productive than expected. The parameter γ measures the degree that co-racial hires are favored in dismissals, where $1+\gamma$ is the proportional increase in the probability that co-racial hires are retained relative to out-group hires.

In combination, co-racial hiring and racial differences in dismissals determine a firm's steady

[15] Managers may also have biased beliefs about group differences in productivity that favor their own group (Lepage, 2021).

state nonwhite share of hires, $\tilde{\pi}$:

$$\tilde{\pi} = \frac{f(\theta, \beta + \gamma + \beta\gamma) \times \theta}{1 - g(\theta, \beta + \gamma + \beta\gamma)}. \tag{4}$$

Critically, the steady state share does not depend on the composition of a firm's employees at any point in time, including at entry.

This model yields two key predictions that we test in the next section. First, firms with white founders—where the initial value of π_{jt} is zero—are more likely to hire white employees than comparable firms with nonwhite founders. Second, for both sets of firms the nonwhite share of hires converges to the same steady state as cumulative hires increase. Though not a necessary consequence of co-racial hiring, we also test whether co-racial hires are less likely to be dismissed.

5 Co-Racial Hiring Evidence

In this section we test for co-racial hiring (Section 5.1) and analogous racial differences in dismissal rates (Section 5.2).

5.1 Hiring

A key prediction of our co-racial hiring model is that firms with white and nonwhite founders are initially more likely to hire co-racial employees, but their hiring behavior converges over the course of the firm life cycle. In other words, a firm's steady state composition does not depend on its initial social conditions, but initial conditions do influence the firm's transitional path to steady state.

To test this prediction, we estimate the following variant of (2), where we allow the $\eta^{s,n}$ coefficients to vary with the race of the firm's founder:

$$\begin{aligned}\log(E(\text{NONWHITE}_{jh}|\cdot)) = &\sum_s \sum_n \sum_r \eta^{s,n,r} \times \mathbb{1}_{\{S(j)=s\}} \times \mathbb{1}_{\{N(j,h)=n\}} \times \mathbb{1}_{\{R(j)=r\}} \\ &+ \tau_{t(j,h)} + \mu_{m(j)} + \omega_{o(j,h)} + \epsilon_{jh},\end{aligned} \tag{5}$$

where $R(j)$ categorizes HQ establishments by founder race. As above, we restrict to a balanced sample of firms for estimation.

We plot the η 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 C.2; the results are similar.) The pattern fits the prediction. For early hires, the racial composition of new hires is tied to founder race. For initial hires, the probability that the hire is nonwhite is 22–32 log points higher at firms with a nonwhite founder compared to firms with a white founder. The gap declines steeply in cumulative hires. By the 50$^{\text{th}}$ hire, the gap declines to about 15 log points, and to about 10 log points by the 200$^{\text{th}}$ hire. After the 400th hire, the gap hovers between 3 and 5 log points. Observably similar firms with white and nonwhite founders appear to converge to different workforce compositions, but differences in steady states are small compared to differences in initial hiring.[16]

In Section 3 we noted that, holding cumulative hires fixed, the probability that the hire is nonwhite is increasing in the total number of hires the firm will make. Interestingly, this is true for both firms with white and nonwhite founders, suggesting that co-racial hiring cannot explain this pattern. We leave further exploration of this pattern to future research.

We conduct an analogous exercise for new establishments that are subsidiaries of preexisting firms.[17] We characterize these establishments by the racial composition of the firm's incumbent employees (at preexisting establishments) when the new establishment first appears in the RAIS data. We divide establishments into two bins by nonwhite share of incumbent employees: 0–50% and 51%–100%. We estimate a model analogous to (5) that divides establishments into these two categories.

We plot η coefficient estimates in Panel B of Figure 4. The findings are similar to what we observe for new firms. 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 diminish as the establishment's cumulative hires increase.

We focus on the convergence between firms with white and nonwhite founders (or majority white and nonwhite incumbent employees), but this convergence pattern is more general: the dispersion in nonwhite share of hires across firms decreases as cumulative hires increase.

We extend (5) and estimate the following model:

$$\log(E(\text{NONWHITE}_{jh}|\cdot)) = \tau_{t(j,h)} + \omega_{o(j,h)} + \theta_{jN(j,t)} + \epsilon_{jh} \tag{6}$$

where $\theta_{jN(j,t)}$ are firm fixed effects for bin of hires N: 1–50, 51–100, ..., 451–500. We estimate the model using a balanced panel of firms, this time combining both entrant firms and new establishments from preexisting firms. We standardize the θ estimates to be mean zero with standard deviation one across firms for the 1–50 bin within each firm category (defined by total observed hires).

Figure 5 plots the standard deviation of θ estimates by cumulative hires. We find that the dispersion of firm effects decreases in cumulative hires. For example, for firms that we observe making at least 500 hires, the standard deviation of firm effects decreases by 8% from the first bin of hires (1-50) to the last bin (451-500).

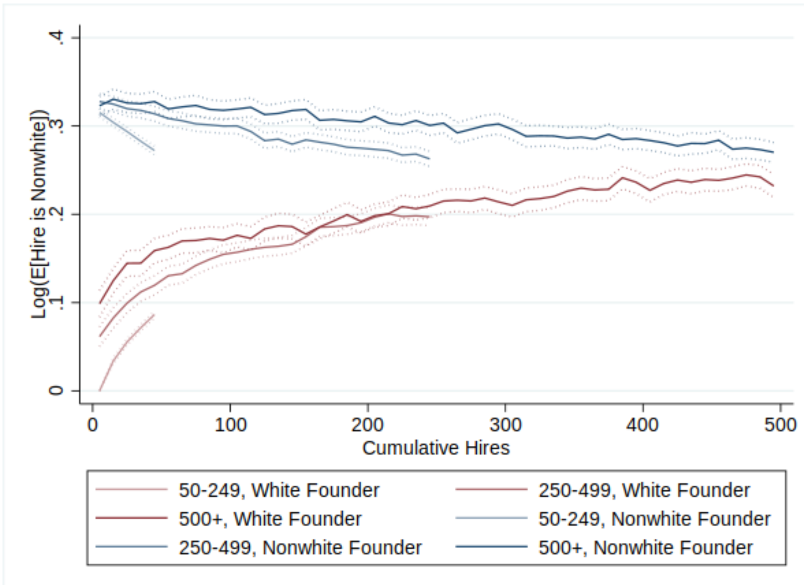
We next test another model prediction: conditioning on π_{jh}, the nonwhite share of incumbent employees at the firm, should attenuate the relationship between a firm's cumulative hires to date and its nonwhite share of hires. That's because this relationship is mediated by π_{jh} via co-racial

[16] As in Section 3, the patterns depicted in Figure 4 are insensitive to the job characteristics we condition on (see Appendix C and Appendix Table C.7).

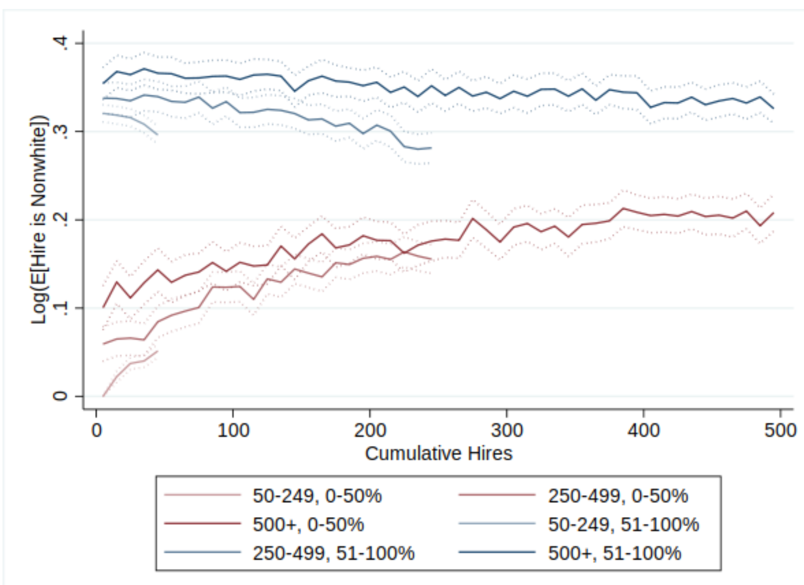
[17] For descriptive statistics on new subsidiary establishments and their hires, see Appendix Table C.4 and Appendix Table C.5.

Figure 4
Founder Race and Convergence in Nonwhite Share of Hires

(a) Entrant Firms

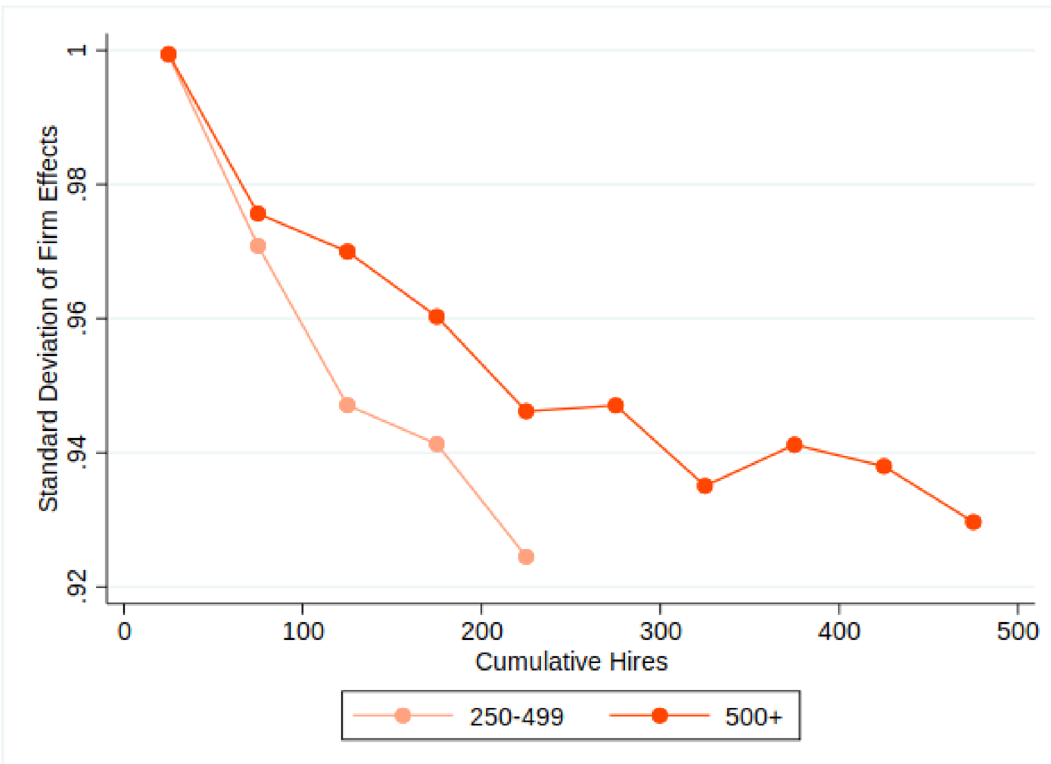


(b) New Subsidiary Establishments



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 $\eta^{s,n,r}$ coefficient estimates from equation (5), summarizing the relationship between a firm's racial composition of hires, its cumulative hires to date (n), and the race of its founder (r) for each firm category s. Panel B does the same for new establishments of preexisting firms. In Panel A founder race is inferred from the race of the top-paid manager or employee at entry. We exclude the inferred founder from the new hires we consider and when measuring cumulative hires. The omitted category in Panel A is the first ten hires for firms with white founders and 50–249 total observed hires. In Panel B, we classify new establishments into two groups based on the nonwhite share of the parent firm's incumbent employees at the time the new establishment opens: 0–50% non-white, and 51–100% non-white. For this model, the omitted category is the first ten hires for those new establishments with 50–249 total observed hires and where the nonwhite share of incumbent employees is 0–50%.

FIGURE 5
DISPERSION OF FIRM EFFECTS DECREASES IN CUMULATIVE HIRES



Note: This figure illustrates how the dispersion across firms in their nonwhite share of hires evolves as a function of cumulative hires. Firm effects for the nonwhite share of hires, estimated separately for bins of 50 hires, are constructed as described in Section 5.1. We standardize the firm effect (θ) estimates to be mean zero with standard deviation one across firms for the 1–50 bin within each firm category (defined by total observed hires).

hiring. We estimate regression models of the form

$$\log(E(\text{NONWHITE}_{jh}|\cdot)) = \delta \log N(j,h) + \tau_{t(j,h)} + \mu_{m(j)} + \omega_{o(j,h)} + \sigma_{S(j)} + f(\pi_{jh}) + \epsilon_{jh} \qquad (7)$$

using the balanced sample of entrant firms, where $f(\pi_{jh})$ is a 10-piece linear spline in π_{jh} and $\sigma_{S(j)}$ are $S(j)$ fixed effects, where $S(j)$ categorizes the firm's total observed hires. δ captures the relationship between cumulative hires ($N(j,h)$) and the nonwhite share of hires. The model predicts that conditioning on π_{jh} will attenuate the estimated δ coefficient.

Table 3 presents δ coefficient estimates. In column 1 we do not condition on π_{jh}. The estimated δ coefficient is 0.0157, indicating that a 10% increase in cumulative hires increases the probability that the next hire is nonwhite by about 0.2%. Conditioning on π_{jh} (column 2) reduces the estimated δ coefficient by over 80%, to 0.0026. This pattern is consistent with co-racial hiring mediating the relationship between firms' cumulative hires and nonwhite share of hires. In columns 3 and 4 we allow both the δ coefficient and $\sigma_{S(j)}$ fixed effects to vary with the race of the founder. For both firms with white and nonwhite founders, conditioning on the nonwhite share of incumbent employees substantially attenuates the relationship between cumulative hires and the nonwhite share of hires.

5.2 Dismissals

We test whether firms with white and nonwhite founders are less likely to dismiss co-racial hires, and whether racial differences in dismissal rates converge between firms with white and nonwhite founders as cumulative hires increase. Using the balanced panel of entrant firms we estimate regression models of the form

$$\log(E(\text{DISMISSED-12M}_{jh}|\cdot)) = \tau_{t(j,h)} + \omega_{o(j,h)} + \psi_{jN(j,h)} + \psi^{NW}_{jN(j,h)} + \epsilon_{jh}, \qquad (8)$$

where DISMISSED-12M$_{jh}$ is an indicator for whether a hire is dismissed within 12 months of their hire date and $\psi_{jN(j,h)}$ are firm by cumulative hire bin fixed effects. Here we group cumulative hires into buckets of 100 hires: 1–100, 101–200, and so on, up to 401–500.[18] $\psi^{NW}_{jN(j,h)}$ are firm by cumulative hire bin by nonwhite fixed effects. Hence, $\psi^{NW}_{jN(j,h)}$ measures the firm-specific racial gap in log twelve-month dismissal rates for a particular set of hires (e.g., hires 1 through 100).

Figure 6 depicts the averages of $\psi^{NW}_{jN(j,h)}$ by cumulative hire bin separately for firms with white and nonwhite founders. For the first 50 hires at firms with white founders, the twelve-month dismissal rate is about 8% higher for nonwhite hires. This declines to a 5% gap for hires 451–500. By contrast, for the first 50 hires at firms with nonwhite founders, the twelve-month dismissal rate is about 4% *lower* for nonwhite hires. There is essentially no racial difference in dismissal rates at firms with nonwhite founders after 250 hires.[19]

More generally, the dispersion in firm-specific dismissal rate differences is decreasing in cumu-

[18] We exclude firms with 50–249 hires and limit to hires 1–200 for firms with 250–499 hires.

[19] The pattern is similar if we include all separations rather than limit to dismissals (see Appendix Figure C.3).

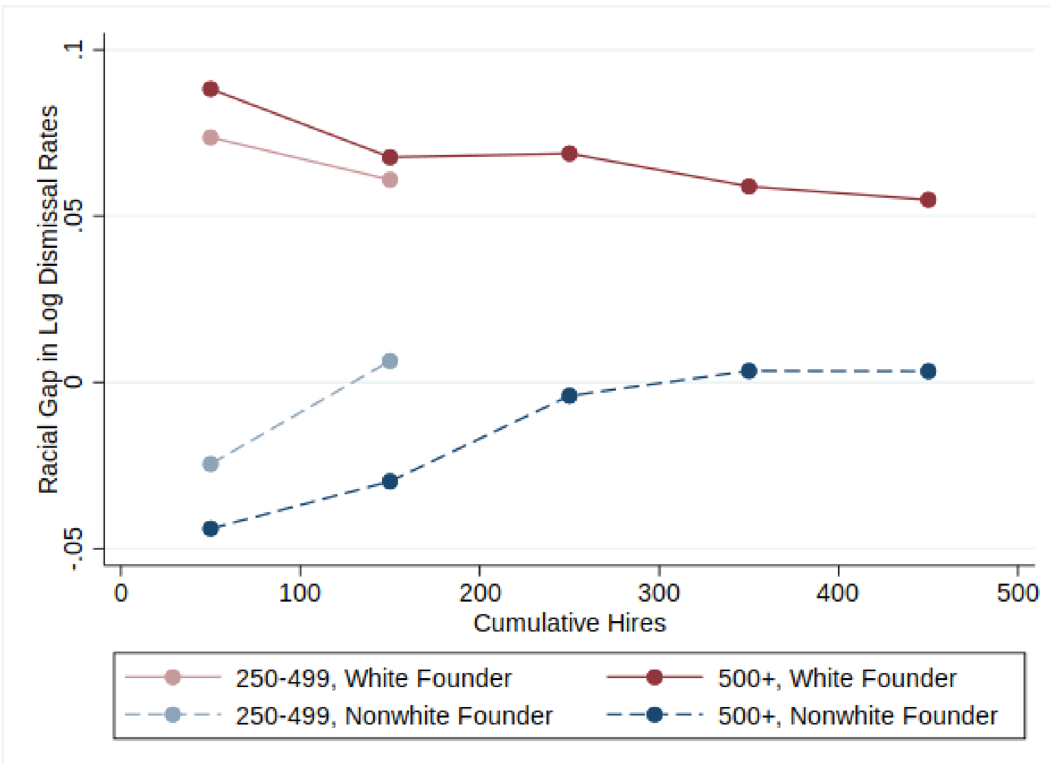
Table 3

Conditioning on Composition of Incumbent Employees Attenuates Life Cycle Pattern

	(1)	(2)	(3)	(4)
log Cumulative Hires (N)	0.0157	0.0026		
	(0.0003)	(0.0004)		
log Cumulative Hires (N)			0.0485	0.0099
$\times$ White Founder			(0.0005)	(0.0005)
log Cumulative Hires (N)			-0.0149	-0.0083
$\times$ Nonwhite Founder			(0.0004)	(0.0004)
NW Share of Employees (π) Spline		✓		✓
Total Hires Bin FEs	✓	✓		
Total Hires Bin $\times$ Founder Race FEs			✓	✓
Occupation FEs	✓	✓	✓	✓
Microregion FEs	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓
N Observations		20,331,798		

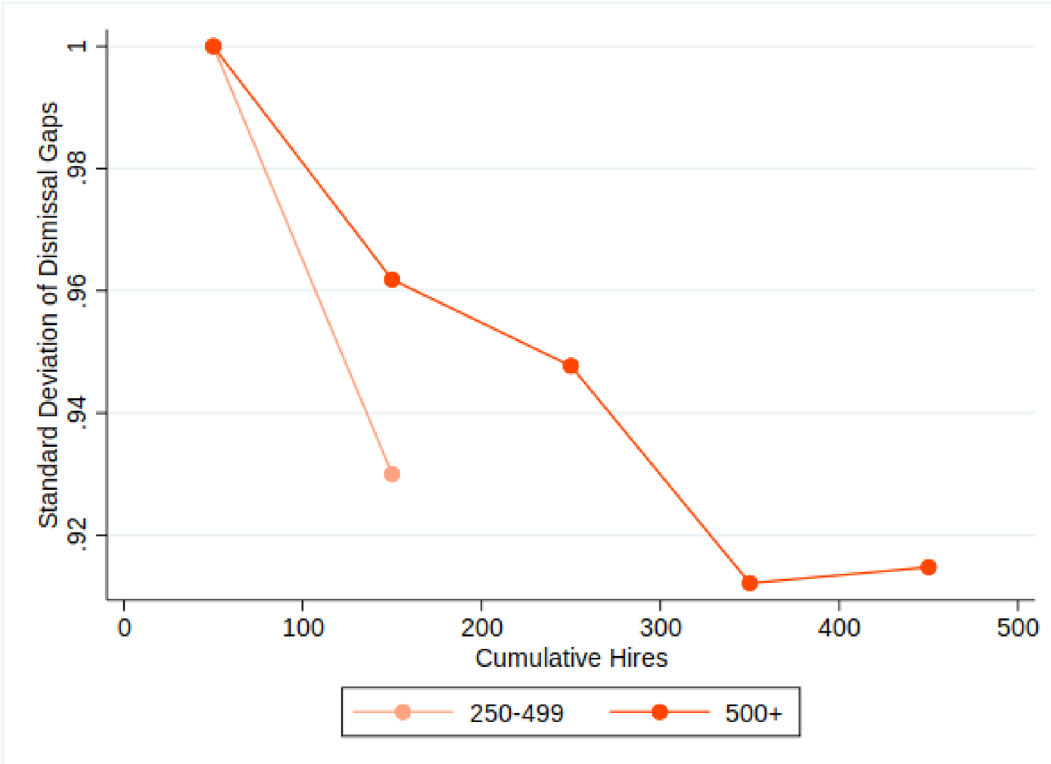
Note: This table presents δ coefficient estimates from equation (7), which summarizes the relationship between the racial composition of a firm's hires and its cumulative hires to date, with and without conditioning on the nonwhite share of the firm's incumbent employees. The outcome is an indicator for whether a hire is nonwhite. All specifications include year, microregion, and occupation fixed effects. Columns 2 and 4 include 10-piece linear splines in the nonwhite share of firm's employees in the month prior to the hire. Columns 1 and 2 include fixed effects for $S(j)$, which categorizes the firm's total observed hires (50–249, 250–499, and 500 or more). Columns 3 and 4 allow those fixed effects to vary with the race of the founder. Founder race is inferred from the race of the top-paid manager or employee at entry. The model is estimated via Poisson quasi maximum likelihood.

Figure 6
Convergence in Dismissal Rates



Note: This figure plots the adjusted, firm-level nonwhite-white difference in log 12-month dismissal rates (ψ_{jN}^{NW}) as a function of founder race and cumulative hires. Firm-specific racial differences in dismissal rates, which can vary with cumulative hires, are constructed as described in equation (8). The model is estimated via Poisson quasi maximum likelihood. Cumulative hires are divided into buckets of 100 hires: 1–100, 101–200, and so on, up to 401–500. The estimation sample is limited to hires 1–200 for firms with 250–499 hires and hires 1–500 for firms with at least 500 hires. Founder race is inferred from the race of the top-paid manager or employee at entry.

Figure 7
Dispersion of Firm-Specific Racial Differences in Dismissal Rates



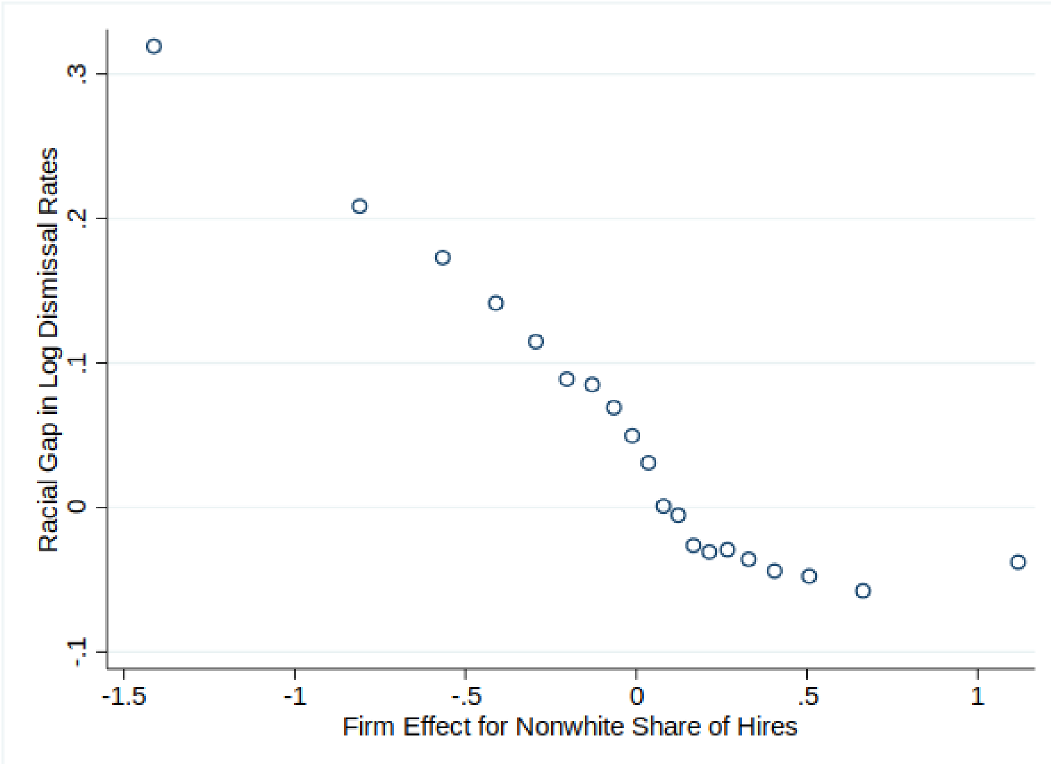
Note: This figure illustrates how the dispersion across firms in their nonwhite-white difference in log 12-month dismissal rates among new hires (ψ_{jN}^{NW}) evolves as a function of cumulative hires, N. Firm-specific racial differences in dismissals rates are constructed as in equation (8). Cumulative hires are divided into buckets of 100 hires: 1–100, 101–200, and so on, up to 401–500. The estimation sample is limited to hires 1–200 for firms with 250–499 hires and hires 1–500 for firms with at least 500 hires. We standardize firm-specific racial differences in dismissals to be mean zero with standard deviation one across firms for the 1–100 bin within each firm category (defined by total observed hires).

lative hires. Figure 7 plots the standard deviation of $\psi_{jN(j,h)}^{NW}$ against cumulative hires, where we standardize firm-specific racial differences in dismissals to be mean zero with standard deviation one across firms for the 1–100 bin within each firm category. For firms that we observe making at least 500 hires, the dispersion of firm-specific dismissal rate differences decreases by 8% from the first bin of hires (1-100) to the last bin (401-500).

The racial composition of hires and racial differences in dismissal rates follow similar patterns: both differences between white-founded firms and nonwhite-founded firms and overall dispersion decrease as cumulative hires increase. This suggests that the composition of hires and dismissals are linked at the firm-level: firms whose hires are disproportionately white have larger (more positive) nonwhite-white differences in dismissal rates. We test this idea.

We re-estimate a modified version of equation (6) that includes microregion fixed effects and uses bins of 100 hires rather than 50 hires so that our estimates of firm effects in racial composition and firm effects in racial differences in dismissal rates are aligned. We focus on firm effects for the

Figure 8
Firms that Hire More Nonwhites Are Less Likely to Dismiss Them



Note: This figure is a binscatter that depicts the relationship between firms' (adjusted) nonwhite share of hires and their (adjusted) nonwhite-white difference in log dismissal rates, limiting to firms' first 100 hires. Firm effects for nonwhite share of hires are estimated as in equation (6), except that the model includes microregion fixed effects and uses bins of 100 hires rather than 50 hires. Firm-specific racial differences in log dismissal rates are estimated as in equation (8). Both models are estimated via Poisson quasi maximum likelihood. Firms are grouped into ventiles by $\theta_{j[1-100]}$, their firm effect for nonwhite share of their first 100 hires.

first 100 hires, $\theta_{j[1-100]}$ and $\psi^{NW}_{j[1-100]}$. We then relate $\theta_{j[1-100]}$ and $\psi^{NW}_{j[1-100]}$ estimates at the firm level.

Figure 8 plots the results using a binscatter. Indeed, there is a clear negative relationship between $\theta_{j[1-100]}$ and $\psi^{NW}_{j[1-100]}$. Firms whose hires are disproportionately white tend to have large and positive nonwhite-white differences in dismissal rates, while firms whose hires are disproportionately nonwhite tend to have small or negative nonwhite-white differences in dismissal rates.

6 Referral Hiring Evidence

Co-racial hiring can take several forms. Quantifying the relative importance of each form is beyond the scope of this paper. Instead, we provide evidence that referral hiring is one empirically relevant form of co-racial hiring in our context. We show that referral networks are racially segregated and socially-connected hires are less likely to be dismissed.

6.1 Empirical Model of Referral Effects

While there is considerable evidence that referrals are an important driver of hiring and job search (see Ioannides and Loury (2004) and Topa (2019) for reviews), none of it covers Brazil. If referrals are important in our setting, we should observe that firms are more likely to hire job seekers when they have a connection to one of their incumbent employees. Likewise, job seekers should be more likely to move into firms where they have a social connection. To check whether this is in fact the case, we adopt a now-common approach, which is to use relational information contained in our administrative data to proxy for the presence of a true social contact. In our case, we are able to observe whether any two workers have been employed in the same job in the past.

Following Eliason et al. (2020), we evaluate the importance of referral hiring by modeling data with the following dyadic structure. Each observation pairs a worker, i, who leaves employer j and may be hired by potential destination employer k in the following year. The binary outcome variable, P_{ijk}, takes value 1 if worker i moves from j to k. The variable of interest $C_{ijk} = 1$ if i has a social connection to another worker (in our case, a past coworker) already employed at destination k and $C_{ijk} = 0$ otherwise.[20]

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

$$P_{ijk} = \alpha_{jk} + X_{ij}\beta + \lambda C_{ijk} + \varepsilon_{ijk}. \tag{9}$$

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. The parameter α_{jk} captures unobserved heterogeneity in the probability of hiring between origin-destination pairs, and X_{ij} are observed characteristics of the worker and the origin employer. We include education, race, gender, and age in X_{ij}.

6.1.1 Identification of Referral Effects

We use a novel combination of strategies to infer referral effects in hiring from our information on past coworkers. In equation (9), the connection 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. A positive estimate of λ does not necessarily imply that there is a true effect of the coworking relationship itself on hiring. Instead, the employer might learn something new about the origin firm from the incumbent employee that makes them 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 same restaurant later.

[20] We consider all workers who change jobs when constructing dyads for the analysis reported in this section. In Appendix C.5 we report results restricted to transitions involving workers that separate during mass displacement events.

To isolate the effect of the connection between the coworkers, we pursue a complementary strategy. We compare the effect on hiring of true coworker connections with the effects associated with what we call *placebo* coworkers: pairs that have held the same job, but not at the same time. In our data, we observe the exact range of months over which a worker is employed in any job. Using this information, we measure, for all pairs of workers ever employed in the same job, the number of months their employment spells overlap. By construction, the amount of overlap is positive when the two were actually coworkers. When they were not actual coworkers, the measured overlap is negative and captures the number of months that elapsed between their employment spells. We define placebo coworkers as those pairs of workers with overlap between -12 months and 0 months.[21] We can then identify the connection effect on any outcome by comparing dyads involving workers with true coworker links to the potential destination firm and those with placebo links.

6.1.2 Data for the Referral Analysis

We estimate connection effects on hiring using data covering all of Brazil between 2013 and 2017. Table 4 describes the sample, which is composed of 303,338,866 dyads. To construct the dyads, we associate each worker hired in a given year with the plant from which they separated in the prior year (the origin plant) for all hires between 2013–2017. For each origin job, we construct the set of potential destination plants as those to which at least on separating worker from the origin job moves. For each separating worker, we assign one observation for each such potential destination.[22] We code two workers as coworkers if they were employed in the same establishment and the same occupation at the same time.[23] Further details appear in Appendix A.1.3.[24]

Table 4 reports that just 0.082 percent of dyads capture cases where worker i is hired by potential job k. The transitioning worker has a true linked coworker connection in 4.1 percent of dyads, while 8.4 percent of observations have either a true link or a placebo link. 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

[21] The threshold of -12 means that two workers that were employed in the same firm more than one year apart are not counted as placebo coworkers.

[22] 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.

[23] 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.

[24] 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.

Table 4
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
Num. Obs.	303,338,866	1,353,787	9,216,640

Note: Column 1 includes pairs of job changers matched to potential destinations. Column 3 describes the population of incumbent workers who are linked to some hired worker via a past coworking relationship.
Source: RAIS, 2013–2017.

of all workers in RAIS. Note that the incumbents in column 3 measure characteristics of linked incumbents from true coworker connections only and not from placebo links.

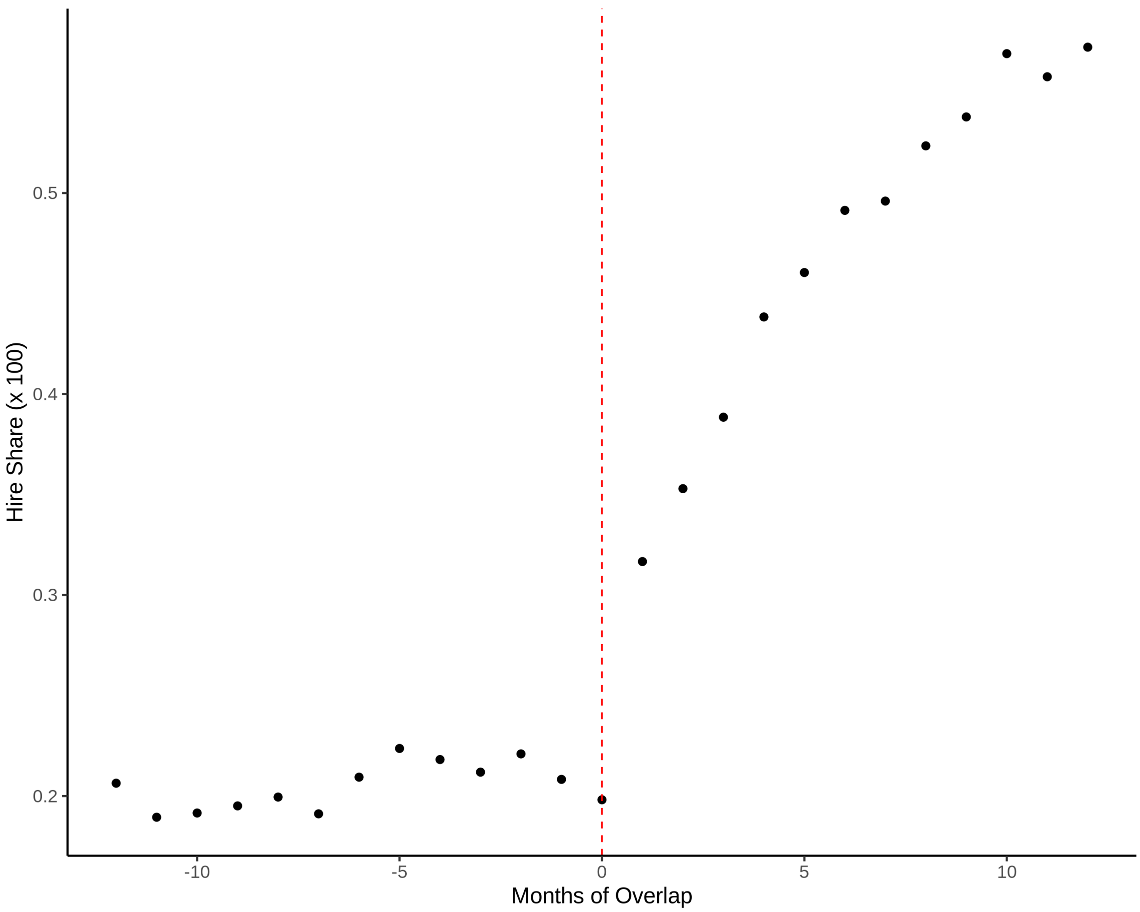
6.1.3 True Coworker Experiences Affect Hiring

Figure 9 illustrates the contrast between true coworkers and placebo coworkers. The horizontal axis measures the number of months of coworker overlap. Each point in the figure measures the share of dyads in which a hire occurred (multiplied by 100) for a given value of overlap. The figure documents three key findings. First, between zero months of overlap and 1 month of overlap we move from dyads where there is not a true coworker connection to one where there is. There we see a discontinuous increase—around 50 percent—in the hire share. This confirms that true coworking relationships have a substantial effect on the likelihood of hiring beyond what can be explained by two workers having similar employment histories. Second, the hire share increases in overlap when overlap is positive, but is not related to overlap when overlap is negative. This is further evidence of a social interaction effect, since the longer two people work together, the more likely they are to know and have useful information about 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). For the placebo connections, the average is double that, around 0.2. This means that the placebo connections do in fact capture information relevant to the hiring outcome.[25]

Consistent with racially segregated social networks, the connection effects we estimate are much stronger for same-race pairs of workers. Figure 10 illustrates. We split the data into four different

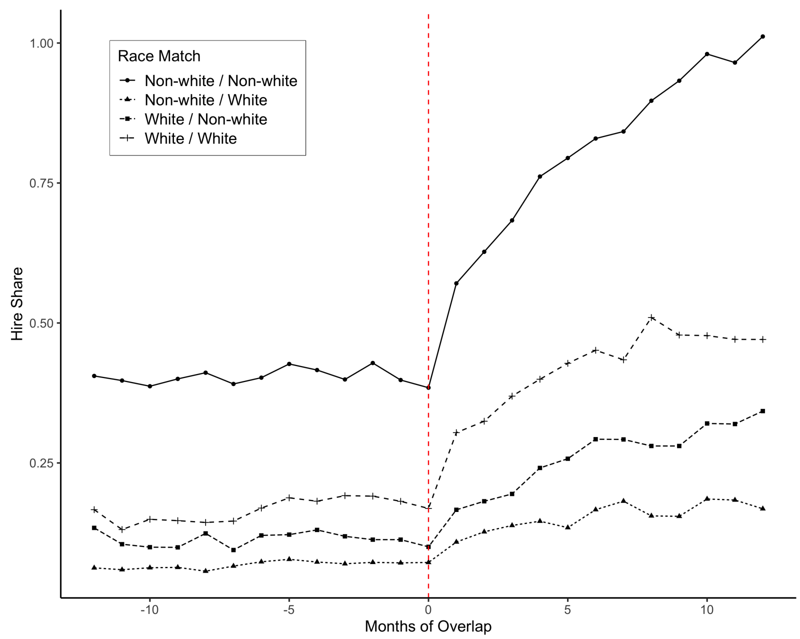
[25] Since our sample is based on all workers who change jobs, one might worry it is biased toward workers with especially productive social connections or toward workers employed in bad matches relative to their outside options. In Section C.5 we show our results are very similar when restricted to workers who separated during mass displacement events.

Figure 9
True Coworker Connections Relative to Placebo Connections



Note: This figure plots the average value of the hiring outcome, P_{ijk} from in equation (9) 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 10
Hiring Share by Coworker Overlap by Incumbent and Job Seeker Race



Note: This figure plots the average value of the hiring outcome in equation (9) 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.

groups based on the racial match of the worker changing jobs and the incumbent worker in the potential destination to whom they have a coworker connection: nonwhite worker linked to a non-white incumbent; nonwhite hire linked to white incumbent, and so on. 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.

6.1.4 Baseline Referral Effects

To estimate the magnitude of connection effects, we extend the Eliason et al. (2020) model (9) to incorporate placebo coworker links:

$$P_{ijk} = \alpha_{jk} + X_{ij}\beta + \lambda C_{ijk} + \lambda^* A_{ijk} + \varepsilon_{ijk}, \tag{10}$$

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

Columns 1, 2, and 3 of Table 5 report estimates of the effect of real coworker links, under different models. Column 1 reports estimates of equation (9), 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 $\lambda = 0.182$, 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 observations with a placebo link. With this specification, $\lambda = 0.222$. 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 equation (10), 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 who have been employed in the same place

TABLE 5
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			

Note: 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.

as one of their incumbents, even when they do not know each other.[26]

Column 4 shows heterogeneity by the match between the race of the job changer and that of the linked incumbent.[27] 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 corroborate the visual evidence in Figure 10, and suggest that referrals are far more common between members of the same 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.

6.2 Referral Hiring and Dismissal Rates

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.

We compare outcomes for connected hires and placebo connected hires as in Section 6.1. For each worker i hired at establishment j, we denote the maximum overlap in prior job spells across combinations of worker i and incumbent workers at establishment j by OVERLAP_{ij}. As discussed in Section 6.1, 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, we set $\text{OVERLAP}_{ij} = -\infty$.

We estimate linear probability models of the form

$$\text{DISMISSED-12M}_{ij} = \sum_{m \in \mathcal{M}} \phi^m \mathbb{1}_{\text{OVERLAP}_{ij} \in m} + \tau_{t(i,j)} + \omega_{o(i,j)} + \psi_j + \epsilon_{ij}, \tag{11}$$

[26] Section C.5 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.

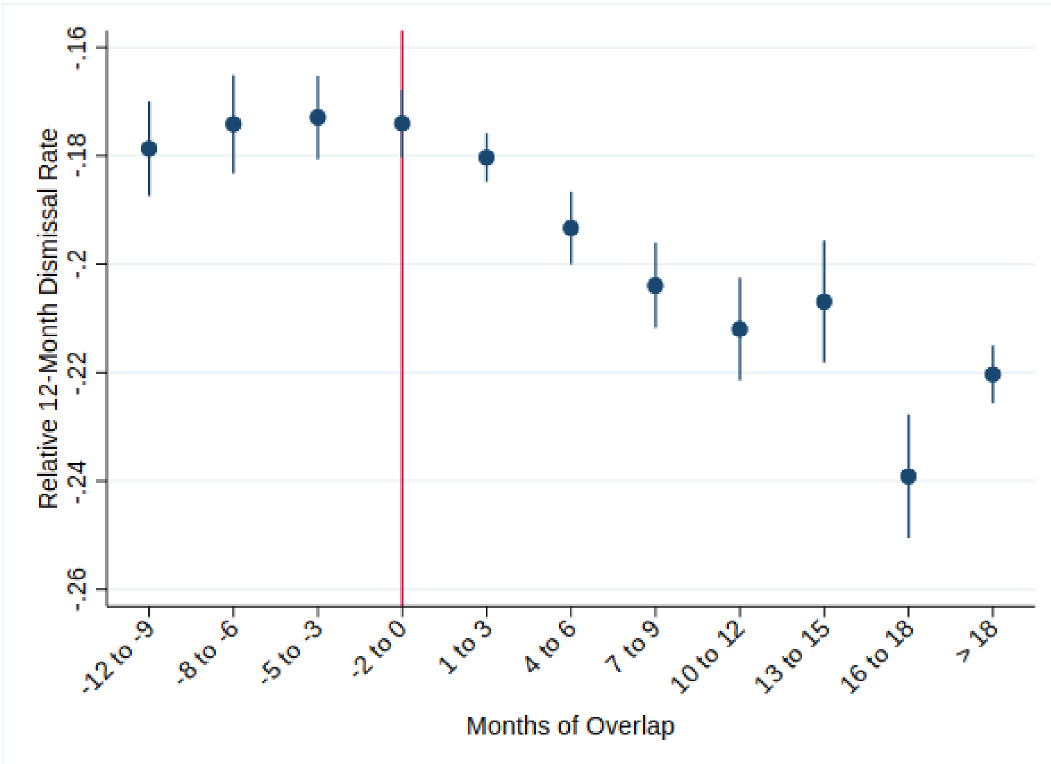
[27] More formally, we estimate an extension of equation (10):

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

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 (10), A_{ijk} indicates the presence of any link, either real or placebo.

Figure 11

Connected Hires Are Less Likely to Be Dismissed



Note: This figure plots the ϕ^m coefficient estimates as described in equation (11). The ϕ^m coefficients identify the dismissal rates of placebo connected and connected hires as a function of overlap relative to hires who are neither.

where $\mathcal{M}$ categorizes OVERLAP$_{ij}$ 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 ϕ^m coefficients identify the dismissal rates of placebo connected and connected hires as a function of overlap relative to hires who are neither.

Figure 11 plots the ϕ^m coefficient estimates. Placebo connected hires are less likely to be dismissed than hires who are neither connected nor placebo connected. Among placebo connected hires, the ϕ^m coefficients hover around -0.18, indicating that placebo connected hires are about 18 percentage points less likely to be dismissed in the first 12 months than peer hires who are neither connected nor placebo connected. For reference, about 32% of hires who are neither connected nor placebo connected are dismissed in the first 12 months. 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. Connected hires with 16 or more months of overlap are about 4 percentage points less likely to be dismissed in the first 12 months than placebo connected hires.

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 pattern as evidence that the referral hires have lower dismissal rates than comparable non-referral hires.

7 Alternative Interpretations

In this section we consider two alternative explanations for our findings that are not forms of co-racial hiring: worker preferences over firms and HR formalization.

7.1 Worker Preferences

We emphasize a demand-side interpretation of the convergence patterns that we document: preferences over workers by race converge between firms with white and nonwhite founders. But there is also a supply side explanation. Rather than firm preferences varying with founder race and cumulative hires, convergence may reflect worker preferences over workplace characteristics. In particular, workers may prefer employers where the founder is of the same race, with this preference weakening as firms' cumulative hires increase.

To evaluate this alternative hypothesis, we build on the insight that worker preferences over employers can be inferred from worker mobility patterns (e.g., Sorkin, 2018; Bagger and Lentz, 2018). We look at two characteristics of new hires: whether they quit their previous job, and whether they were 'poached' from their previous job, which we define as moving from another job without a nonemployment period greater than one month in between job spells. By a revealed preference argument, both behaviors suggest that a new hire preferred their new job over their previous job.

We estimate the following model, separately for white and nonwhite hires:

$$\log(E(Y_{jh}|\cdot)) = \sum_n \sum_r \eta^{n,r} \times \mathbb{1}_{\{N(j,h)=n\}} \times \mathbb{1}_{\{R(j)=r\}} + \tau_{t(j,h)} + \psi_j + \omega_{o(j,h)} + \epsilon_{jh},$$

where Y_{jh} is either an indicator for whether a new hire quit their previous job or an indicator for whether a new hire was poached from their previous employer.

Patterns for $\eta^{n,r}$ coefficients are similar for both outcomes (see Appendix Figure C.5 for details). Workers appear to prefer firms later in the firm life cycle, perhaps after they've become more established. Yet mobility patterns do not differ much by worker race, and there is no interaction between worker and founder race. To the extent that mobility patterns capture worker preferences, there is little evidence that supply-side preferences contribute to the findings documented in Section 5.

7.2 HR Formalization

We next examine whether the hiring dynamics we document are driven by the formalization of HR (Dobbin, 2009). We take advantage of the fact that we can identify HR-related occupations in the RAIS data. We find that the convergence pattern is similar for firms that do and do not hire anyone in an HR position among their first 500 hires (see Appendix Figure C.6). HR formalization, at least as measured by the presence of HR professionals, does not appear to play a significant role for our findings.

8 Discussion

We document and examine a striking pattern: as a firm's cumulative hires increases, so does the probability that its next hire is nonwhite. The pattern is not explained by changes in observable job characteristics.

Instead, we argue that the pattern reflects two labor market features: racial disparities in entrepreneurship and co-racial hiring. Consistent with co-racial hiring, we find that: firms with white and nonwhite founders are more likely to hire white and nonwhite employees and these differences dissipate as firms' cumulative number of hires increases. Dismissals follow an analogous pattern: firms are less likely to dismiss recent hires of the same race as the founder and racial differences in dismissal rates are also decreasing in firms' cumulative number of hires.

We find that, for the average firm, the nonwhite share of hires is about 4% lower than the steady state. This suggests that the factors that lead firms to hire fewer nonwhite workers early in the firm life cycle are reducing relative demand for nonwhite workers in the aggregate. Moreover, our findings can help explain why nonwhite workers are more likely to be dismissed from their jobs and have less seniority relative to their co-workers (see Appendix C for details).

Given data constraints, we are agnostic about the relative importance of different forms of co-racial hiring for the patterns we document and whether co-racial hiring is efficient. We provide evidence that our findings are at least in part driven by referral hiring. In particular, we show that referral networks are racially segregated and socially-connected hires are less likely to be dismissed.

Regardless of the channel, our findings suggest how policy can reduce or contribute to racial inequality in labor demand. First, our findings provide a 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 may have short-run costs—for example, an increase in dismissal rates—but our findings suggest it may lead to persistent reductions to racial inequality in labor demand (Miller, 2017). While affirmative action policies often exclude small firms, our findings suggest these policies may be particularly effective at small (or young) firms.

Second, we demonstrate a direct link between racial differences in entrepreneurship and labor demand, motivating policies that encourage entrepreneurship among underrepresented groups.

Third, 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.

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.

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.

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.

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," 2022. 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.

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, 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.

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.

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

Cornwell, Christopher, 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.

de Geografia e Estatística IBGE, Instituto Brasileiro, "Pesquisa Nacional por Amostra de Domicílios (PNAD) 2003-2017," 2021.

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

Dobbin, Frank, *Inventing Equal Opportunity*, Princeton University Press, 2009.

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 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.

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.

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 Retail 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.

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

Ioannides, Yannis M. and Linda Datcher Loury, "Job information networks, neighborhood effects, and inequality," *Journal of Economic Literature*, 2004, *42*, 1056–1093.

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.

Lepage, Louis-Pierre, "Endogenous Learning, Persistent Employer Biases, and Discrimination," June 2021. Unpublished manuscript.

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.

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

Pan, Jessica, "Gender Segregation in Occupations: The Role of Tipping and Social Interactions," *Journal of Labor Economics*, 2015, *33* (2), 365–408.

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.

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

Rubineau, Brian and Roberto M. Fernandez, "Missing Links: Referrer Behavior and Job Segregation," *Management Science*, November 2013, *59* (11), 2470–2489.

_ and _ , "Tipping Points: The Gender Segregation and Desegregating Effects of Network Recruitment," *Organization Science*, 2015, *26* (6), 1646–1664.

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.

Schelling, Thomas C., "Dynamic Models of Segregation," *Journal of Mathematical Sociology*, 1971, *1*, 143–186.

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

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.

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 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.

Online Appendix: The Dynamic Effects of Co-Racial Hiring

Conrad Miller

Ian Schmutte

January 2023

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.[28]

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

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 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 Model Appendix

We first describe the reduced form implications of the model and then provide a micro foundation.

Let θ denote the nonwhite share of the candidate pool. In the absence of co-racial hiring, with probability p a candidate from either group will send a productivity signal strong enough to be hired. Suppose that, with co-racial hiring, this probability increases to $(1+\beta)p$ for the in-group.[29] If the in-group is nonwhite, the probability that a new hire is nonwhite is

$$\frac{(1+\beta)p\theta}{(1+\beta)p\theta + p(1-\theta)} = \frac{\theta(1+\beta)}{1+\theta\beta}$$

and if the in-group is white, the probability that a new hire is nonwhite is

$$\frac{p\theta}{p\theta + (1+\beta)p(1-\theta)} = \frac{\theta}{1+\beta(1-\theta)}.$$

Let π denote the nonwhite share of the firm's incumbent employees. The probability that the next hire is nonwhite is

$$\begin{aligned} P(\text{NONWHITE}) &= \pi\left(\frac{\theta(1+\beta)}{1+\theta\beta}\right) + (1-\pi)\left(\frac{\theta}{1+\beta(1-\theta)}\right) \\ &= \frac{\theta}{1+\beta(1-\theta)} + \pi\left(\frac{\theta(1+\beta)}{1+\theta\beta} - \frac{\theta}{1+\beta(1-\theta)}\right) \\ &= f(\theta,\beta)\theta + g(\theta,\beta)\pi. \end{aligned}$$

During a probationary period, the true productivity of new hires is revealed. New hires may be dismissed if their productivity is below some threshold. As we show below, co-racial hiring may improve retention rates for in-group hires as well.

Benson et al. (2022) develop a model of hiring that encompasses multiple forms of co-racial hiring: taste-based discrimination, screening discrimination, and production complementarities. We describe the model here.

The firm must fill a vacancy. Among incumbent employees, a randomly selected employee is chosen as the hiring manager. The manager screens applicants. Suppose worker productivity follows $y \sim N(\mu, \sigma_0^2)$. The manager receives a noisy signal for this productivity, $\tilde{y}$, where $\tilde{y} = y + \epsilon$ and the noise $\epsilon \sim (0, \theta_\epsilon^2)$ is independent of y.

The expected productivity of an applicant given the signal is

$$\hat{y} = E[y|\tilde{y}] = \frac{\sigma_0^2}{\sigma^2 + \sigma_\epsilon^2}\tilde{y} + \frac{\sigma_\epsilon^2}{\sigma^2 + \sigma_\epsilon^2}\mu.$$

[29] Alternatively, we can think of co-racial hiring as affecting the composition of candidates that the firm encounters. For example, with referral hiring, the firm may meet a referral candidate with some probability and an external market candidate otherwise.

The estimate $\hat{y}$ is distributed normally with mean μ and estimator variance $\eta^2 = \frac{\sigma_0^4}{\sigma_0^2+\sigma_\epsilon^2}$.

Conditional on $\hat{y}$, the realized ability y is distributed normally with mean $\hat{y}$ and residual variance $\gamma^2 = \frac{\sigma_0^2\sigma_\epsilon^2}{\sigma_0^2+\sigma_\epsilon^2}$.

Assume that managers have some threshold for expected productivity y^*.[30] The hiring probability is

$$p = P(\hat{y} \geq y*)$$

we also assume $y^* \geq \mu$ so at most half of all applicants are hired.

Finally, suppose the firm eventually learns the actual productivity y during a probationary period, and dismisses a worker if it is more than τ below and cutoff y^*. Turnover then equals

$$x = P(y \leq y^* - \tau | \hat{y} \geq y^*)$$

In all three cases below, the hiring probability is higher for in-group candidates than out-group candidates and turnover is lower for in-group hires than out-group hires. See Benson et al. (2022) for proofs.

Taste-Based Discrimination In this case, managers apply a lower hiring threshold for in-group hires (s for same race) than out-group hires (c for cross race), so $y_s^* < y_c^*$. (see Proposition 1 of Benson et al. (2022).)

Screening Discrimination In this case, signals are more precise for in-group applicants than out-group so with signal variance $\sigma_s^2 < \sigma_c^2$. (see Proposition 2 of Benson et al. (2022).)

Complementary Production In this case, workers are more productive under same-race managers, $\mu_s > \mu_c$. (see Proposition 3 of Benson et al. (2022).)

Referral hiring can be interpreted as similar to any of these cases (Topa, 2019): incumbent employees may prefer to work with their existing social connections; they may have more information about their match quality ex-ante; and they may be more productive when working with existing social connections. In addition, referral hiring may change the racial composition of the applicant pool.

[30] For simplicity, we assume this threshold is fixed no matter the identity of the hiring manager. A natural extension would be to allow this threshold to depend on the hiring manager's group, which would affect the size of the in-group.

C Additional Exhibits

C.1 Estimating Slope of Relationship between Nonwhite Share and Cumulative Hires

We estimate a more parametric variant of equation (1) where we replace indicator variables for cumulative hire bins with $\log N(j,h)$, or log cumulative hires to date. This summarizes the relationship between the nonwhite share of a firm's hires and its cumulative hires to date using a single elasticity. We estimate

$$\log(E(\text{NONWHITE}_{jh}|\cdot)) = \delta \log N(j,h) + \tau_{t(j,h)} + \psi_j + X_{jh} + \epsilon_{jh} \tag{C.1}$$

where NONWHITE_{jh} is an indicator for whether hire h at firm j is nonwhite, $N(j,h)$ denotes cumulative hires to date at firm j, ψ_j are firm fixed effects, $\tau_{t(j,h)}$ are year fixed effects, and X_{jh} is a vector of additional controls for job characteristics. We vary this set of controls across specifications. We limit the estimation sample to firms' first 500 hires. δ captures the relationship between cumulative hires ($N(j,h)$) and the nonwhite share of hires.

Table C.7 presents δ slope estimates. Column 1 includes only year fixed effects and firm fixed effects as controls and yields a δ estimate of 0.0152. As reflected in Figure 2, including occupation fixed effects (column 2) attenuates the δ coefficient to 0.0114. Replacing occupation fixed effects with firm by occupation fixed effects (column 3) or firm by occupation by contemporaneous firm size fixed effects has little effect.

Columns 5 through 8 are analogous to columns 1 through 4 except we allow the δ coefficient to vary with founder race. Column 5 includes only year fixed effects and firm fixed effects as controls. For white-founder firms, the δ coefficient is 0.0497. For nonwhite-founder firms, the δ coefficient is -0.0157. This contrast persists no matter what job characteristics we condition on.

C.2 Linear Probability Models

We repeat the main analyses from the paper using linear probability models (LPM) rather than Poisson models.

The LPM analog to equation (1), which describes the relationship between NONWHITE_{jh} and cumulative hires for all firms pooled to together, that we estimate is

$$\text{NONWHITE}_{jh} = \sum_n \eta^n \times \mathbb{1}_{\{N(j,h)=n\}} + \gamma_{t(j,h)m(j)o(j,h)} + \psi_j + \epsilon_{jh} \tag{C.2}$$

where $\gamma_{t(j,h)m(j)o(j,h)}$ are year by microregion by occupation fixed effects. As in Section 3 we estimate several variants of this model to examine how the η^n coefficients change when we include increasing granular controls for job characteristics. The first specification includes year by microregion fixed effects and firm fixed effects. The second specification includes year by microregion by occupation fixed effects and firm fixed effects. The third specification includes year by microregion fixed effects and firm by occupation fixed effects. The fourth and final specification includes year by microregion fixed effects and firm by occupation by contemporaneous firm size category fixed effects.

The η^n coefficient estimates for each specification are plotted in Figure C.7.

We also expand equation (C.2) (the second specification) and allow the η^n coefficient to depend on founder race. The η^n coefficient estimates are plotted in Panel A of Figure C.8.

Finally, we estimate an analog of equation (5) where we limit the estimation sample to a balanced panel of firms and allow the relationship between NONWHITE_{jh} and cumulative hires to depend on both founder race and total hires:

$$\begin{aligned}\text{NONWHITE}_{jh} = &\sum_s \sum_n \sum_r \eta^{s,n,r} \times \mathbb{1}_{\{S(j)=s\}} \times \mathbb{1}_{\{N(j,h)=n\}} \times \mathbb{1}_{\{R(j)=r\}} \\ &+ \gamma_{tm(j)o(j,h)} + \epsilon_{jh}.\end{aligned} \tag{C.3}$$

The η^n coefficient estimates are plotted in Panel B of Figure C.8.

To examine racial differences in dismissal rates, we estimate

$$\text{DISMISSED-12M}_{jh} = \tau_{t(j,h)} + \omega_{o(j,h)} + \psi_{jN(j,h)} + \psi^{NW}_{jN(J,h)} + \epsilon_{jh}, \tag{C.4}$$

Figure C.9 depicts the averages of $\psi^{NW}_{jN(j,t)}$ by cumulative hire bin separately for firms with white and nonwhite founders.

C.3 Wages

We estimate the relationship between cumulative hires and firm-specific racial differences in wages, analogous to our analysis of dismissal rates in Section 5.2.

$$\log(\text{WAGE}_{jh}) = \tau_{t(j,h)} + \omega_{o(j,h)} + \psi_{jN(j,h)} + \psi^{NW}_{jN(j,h)} + \epsilon_{jh} \tag{C.5}$$

where WAGE_{jh} is starting hourly wages.

Figure C.4 is analogous to Figure 6. Panel A looks at starting wages while Panel B looks at starting wages in the new hires prior job. Consistent with Dias and Rocha (2021), we find that racial wage gaps are smaller in firms with nonwhite founders. For both firms with white and nonwhite founders, the wage gap is declining in cumulative hires. The pattern is similar for prior wages, suggesting that the pattern in Panel A may simply reflect worker selection.

C.4 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

$$\log r_{ijt} \equiv \log n_{jt} - \log q_{ijt}. \tag{C.6}$$

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

$$\log r_{ijt} = \tau_t + \omega_{o(i,t)} + \beta \text{NONWHITE}_i + \nu \log n_{jt} + \epsilon_{it}. \tag{C.7}$$

Estimates for equation (C.7) are presented in Table C.10. 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.

C.5 Analysis of Displaced Workers

Our main estimates of referral effects in Section 6.1.2 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.[31] 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.

Table C.8 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 C.9 reports the same models as Table 5 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

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

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.

Table C.2
Characteristics of Hires at Entrant Firms

		White Founder			Nonwhite Founder		
Hires:	All	All	White	Nonwhite	All	White	Nonwhite
	(1)	(4)	(5)	(6)	(7)	(8)	(9)
Nonwhite (%)	38.1	27.2	0.0	100.0	62.0	0.0	100.0
Log Wage	1.784	1.816	1.835	1.765	1.715	1.777	1.681
		(0.433)	(0.442)	(0.402)	(0.396)	(0.421)	(0.376)
Male (%)	65.4	64.3	62.8	68.1	68.6	66.4	71.1
Age	30.8	30.7	30.8	30.3	31.0	31.4	30.8
$<$ HS	23.1	22.3	21.2	25.4	24.9	22.4	26.4
HS Grad	69.0	69.0	69.1	68.6	69.0	69.3	68.9
College Grad	7.9	8.7	9.8	5.9	6.1	8.3	4.8
12M Separation	57.3	57.5	57.0	59.1	56.7	57.3	56.3
12M Dismissal	39.7	39.0	37.8	42.0	41.4	40.2	42.2
J-J Move	66.4	67.4	68.1	65.4	64.1	65.8	63.0
Quit Prior Job	18.7	20.4	21.4	17.6	15.0	17.0	13.7
N Hires	64m	44m	32m	12m	20m	8m	12m

This table reports summary statistics for hires at entrant firms in *Relação Anual de Informações Sociais* (RAIS) data for the years 2003–2017. We limit the sample to private sector, indeterminate-length contracts. Each observation is a worker-firm job spell. Entrant firms are identified as described in Section 2.2. Columns 1–3 limit to entrant firms with white founders and columns 4–6 limit to entrant firms with nonwhite founders. Founder race is inferred from the race of the top-paid manager or employee at entry. We compute an hourly wage by deflating average monthly earnings by the product of contracted weekly hours and average weeks per month. Wages refer to starting wages for the job spell.

Table C.1

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 (%)	36.3	0.0	100.0	38.7	0.0	100.0	39.6	0.0	100.0
Log Wage	2.002	2.075	1.874	1.845	1.892	1.770	1.823	1.860	1.766
	(0.676)	(0.715)	(0.581)	(0.554)	(0.584)	(0.495)	(0.471)	(0.489)	(0.435)
Male (%)	66.4	64.3	70.1	67.4	64.8	71.5	66.4	63.8	70.3
Age	33.8	34.1	33.2	30.8	30.9	30.6	31.0	31.2	30.7
< HS	30.9	29.2	34.0	29.1	26.8	32.8	24.2	22.3	27.1
HS Grad	56.8	55.9	58.4	61.1	61.2	61.0	67.2	67.2	67.0
College Grad	12.3	14.9	7.7	9.8	12.0	6.2	8.7	10.5	5.9
Number of Worker-Year Obs.	697m	444m	253m	258m	158m	100m	93m	56m	37m

This table reports summary statistics from the *Relação Anual de Informações Sociais* (RAIS) data for the years 2003–2017. We limit the sample to private sector, indeterminate-length contracts. 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 2.2. We compute an hourly wage by deflating average monthly earnings by the product of contracted weekly hours and average weeks per month.

Table C.3
Characteristics of Hires at Entrant Firms, Balanced Panel

Hires:	All	White Founder: All	White Founder: White	White Founder: Nonwhite	Nonwhite Founder: All	Nonwhite Founder: White	Nonwhite Founder: Nonwhite
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Nonwhite (%)	39.6	29.0	0.0	100.0	63.0	0.0	100.0
Log Wage	1.795	1.824	1.843	1.777	1.730	1.787	1.697
	(0.461)	(0.470)	(0.481)	(0.437)	(0.433)	(0.460)	(0.413)
Male (%)	69.5	68.3	66.7	72.1	73.3	70.9	75.9
Age	30.9	30.8	30.9	30.6	31.2	31.5	31.1
$<$ HS	27.9	26.9	25.5	30.4	30.1	27.2	31.9
HS Grad	65.0	65.2	65.5	64.4	64.5	65.3	64.0
College Grad	7.1	7.9	9.0	5.2	5.4	7.5	4.1
12M Separation	61.2	61.1	60.4	62.9	61.2	62.0	60.8
12M Dismissal	42.4	41.3	39.9	44.8	44.7	43.3	45.5
J-J Move	65.0	65.8	66.5	64.0	63.2	64.6	62.4
Quit Prior Job	19.4	21.1	22.3	18.1	15.6	17.8	14.4
N Hires	23m	16m	11m	5m	7m	3m	4m

This table reports summary statistics for hires at a balanced panel of entrant firms in *Relação Anual de Informações Sociais* (RAIS) data for the years 2003–2017. We limit the sample to private sector, indeterminate-length contracts. Each observation is a worker-firm job spell. Entrant firms are identified as described in Section 2.2. 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 or more total observed hires. Columns 1–3 limit to entrant firms with white founders and columns 4–6 limit to entrant firms with nonwhite founders. Founder race is inferred from the race of the top-paid manager or employee at entry. We compute an hourly wage by deflating average monthly earnings by the product of contracted weekly hours and average weeks per month. Wages refer to starting wages for the job spell.

TABLE C.4

CHARACTERISTICS OF NEW SUBSIDIARY ESTABLISHMENTS

	Pooled	< 50% Nonwhite Incumbents (%)	≥ 50% Nonwhite Incumbents (%)
	(1)	(2)	(3)
≥ 50% Nonwhite Incumbents (%)	24.9	0.0	100.0
Total Hires			
1-19	84.9	84.7	85.4
20-49	8.3	8.4	7.7
50-249	5.8	5.8	5.6
250-499	0.6	0.6	0.7
500-999	0.5	0.4	0.6
1000+	0.1	0.1	0.2
Survival			
After 3 Years	31.3	32.4	27.9
After 5 Years	20.0	21.2	16.5
Industry (%)			
Manufacturing	1.7	1.8	1.3
Construction	21.1	21.5	19.7
Commerce	13.0	12.8	13.6
Transport, Storage, and Mail3.0	3.2	2.7	
Accommodation and Meals	1.6	1.5	1.9
Professional Activities	3.8	3.9	3.4
Administrative Activities	1.8	1.9	1.7
Health and Social Services	9.3	9.3	9.4
Other	44.7	44.1	46.3
Number of Firms	863k	648k	215k

This table reports summary statistics for new establishments in existing firms in the *Relação Anual de Informações Sociais* (RAIS) data for the years 2003–2017. We divide establishments into three bins by nonwhite share of incumbent employees: 0-33%, 34-66%, and 67%-100%.

Table C.5
Characteristics of Hires at New Subsidiary Establishments

		< 50% Nonwhite Incumbents (%)			≥ 50% Nonwhite Incumbents (%)		
Hires:	All	All	White	Nonwhite	All	White	Nonwhite
	(1)	(4)	(5)	(6)	(7)	(8)	(9)
Nonwhite (%)	40.3	32.0	0.0	100.0	64.4	0.0	100.0
Log Wage	1.925	1.960	2.010	1.854	1.822	1.879	1.791
	(0.582)	(0.600)	(0.628)	(0.522)	(0.513)	(0.558)	(0.483)
Male (%)	64.6	63.6	62.8	65.5	68.8	70.4	67.4
Age	31.2	31.2	31.4	30.8	31.2	32.5	30.8
< HS	28.2	27.8	24.8	34.3	29.3	30.0	28.9
HS Grad	60.6	59.7	60.6	57.9	63.3	60.2	65.1
College Grad	11.2	12.5	14.7	7.8	7.4	9.8	6.0
12M Separation	56.1	56.4	55.8	57.8	55.4	56.3	55.0
12M Dismissal	34.1	33.6	32.1	36.6	35.4	35.9	35.1
J-J Move	59.6	60.2	60.5	59.6	57.6	61.7	55.4
Quit Prior Job	30.7	31.8	33.6	27.9	27.8	27.1	28.1
N Hires	14m	11m 7.2m	3.4m	4m	1.3m	2.4m	

This table reports summary statistics for hires at new subsidiary establishments in *Relação Anual de Informações Sociais* (RAIS) data for the years 2003–2017. We limit the sample to private sector, indeterminate-length contracts. Each observation is a worker-firm job spell. Entrant firms are identified as described in Section 2.2. Columns 1–3 limit to new establishment subsidiaries where less than 50% of incumbent firm employees are nonwhite and columns 4–6 limit to new establishment subsidiaries where 50% or more of incumbent firm employees are nonwhite. Founder race is inferred from the race of the top-paid manager or employee at entry. We compute an hourly wage by deflating average monthly earnings by the product of contracted weekly hours and average weeks per month. Wages refer to starting wages for the job spell.

Table C.6
Characteristics of Hires at New Subsidiary Establishments, Balanced Panel

		< 50% Nonwhite Incumbents (%)			≥ 50% Nonwhite Incumbents (%)		
Hires:	All	All	White	Nonwhite	All	White	Nonwhite
	(1)	(4)	(5)	(6)	(7)	(8)	(9)
Nonwhite (%)	42.2	32.2	0.0	100.0	68.3	0.0	100.0
Log Wage	1.973	2.011	2.065	1.899	1.872	1.964	1.830
	(0.614)	(0.631)	(0.661)	(0.546)	(0.555)	(0.626)	(0.514)
Male (%)	63.1	62.1	61.5	63.4	66.9	68.2	65.9
Age	30.4	30.3	30.5	29.9	30.5	31.4	30.0
< HS	26.0	25.6	22.7	31.5	27.3	25.9	27.9
HS Grad	62.5	61.6	62.0	60.6	65.0	62.6	66.2
College Grad	11.4	12.9	15.2	7.9	7.7	11.5	5.9
12M Separation	57.2	57.4	56.6	59.0	56.6	58.5	55.7
12M Dismissal	33.0	32.4	30.6	36.1	34.7	34.9	34.6
J-J Move	55.7	56.6	57.0	55.9	53.4	56.5	51.9
Quit Prior Job	34.7	35.8	37.8	31.6	31.9	33.0	31.4
N Hires	6m	4.2m 2.9m	1.4m	1.6m	0.5m	1.1m	

This table reports summary statistics for hires at a balanced panel of new subsidiary establishments in *Relação Anual de Informações Sociais* (RAIS) data for the years 2003–2017. We limit the sample to private sector, indeterminate-length contracts. Each observation is a worker-firm job spell. Entrant firms are identified as described in Section 2.2. 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 or more total observed hires. Columns 1–3 limit to new establishment subsidiaries where less than 50% of incumbent firm employees are nonwhite and columns 4–6 limit to new establishment subsidiaries where 50% or more of incumbent firm employees are nonwhite. Founder race is inferred from the race of the top-paid manager or employee at entry. We compute an hourly wage by deflating average monthly earnings by the product of contracted weekly hours and average weeks per month. Wages refer to starting wages for the job spell.

TABLE C.7

SLOPE ESTIMATES FOR RELATIONSHIP BETWEEN NONWHITE SHARE OF CUMULATIVE HIRES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log Cumulative Hires (N)	0.0152	0.0114	0.0110	0.0106				
	(0.0002)	(0.0002)	(0.0002)	(0.0003)				
log Cumulative Hires (N)					0.0497	0.0453	0.0434	0.0.0392
$\times$ White Founder					(0.0003)	(0.0003)	(0.0003)	(0.0004)
log Cumulative Hires (N)					-0.0157	-0.01888	-0.0180	-0.0154
$\times$ Nonwhite Founder					(0.0002)	(0.0002)	(0.0002)	(0.0003)
Year FEs	✓	✓	✓	✓	✓	✓	✓	✓
Firm FEs	✓	✓	✓	✓	✓	✓	✓	✓
Occupation FEs		✓				✓		
Firm by Occ. FEs	✓		✓		✓		✓	
Firm by Occ. by Size FEs				✓				✓
N Observations	58,513,055							

Note: This table presents δ coefficient estimates from equation (C.1), which summarizes the relationship between the racial composition of a firm's hires and its cumulative hires to date. The outcome is an indicator for whether a hire is nonwhite. All specifications include year and firm fixed effects. Columns 2 and 6 include 6-digit occupation fixed effects. Columns 3 and 7 include firm by occupation fixed effects. Columns 4 and 8 include firm by occupation by contemporaneous firm size fixed effects. Founder race is inferred from the race of the top-paid manager or employee at entry. The model is estimated via Poisson quasi maximum likelihood.

Table C.8
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

Note: 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.
Source: RAIS, 2013–2017.

Table C.9
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.

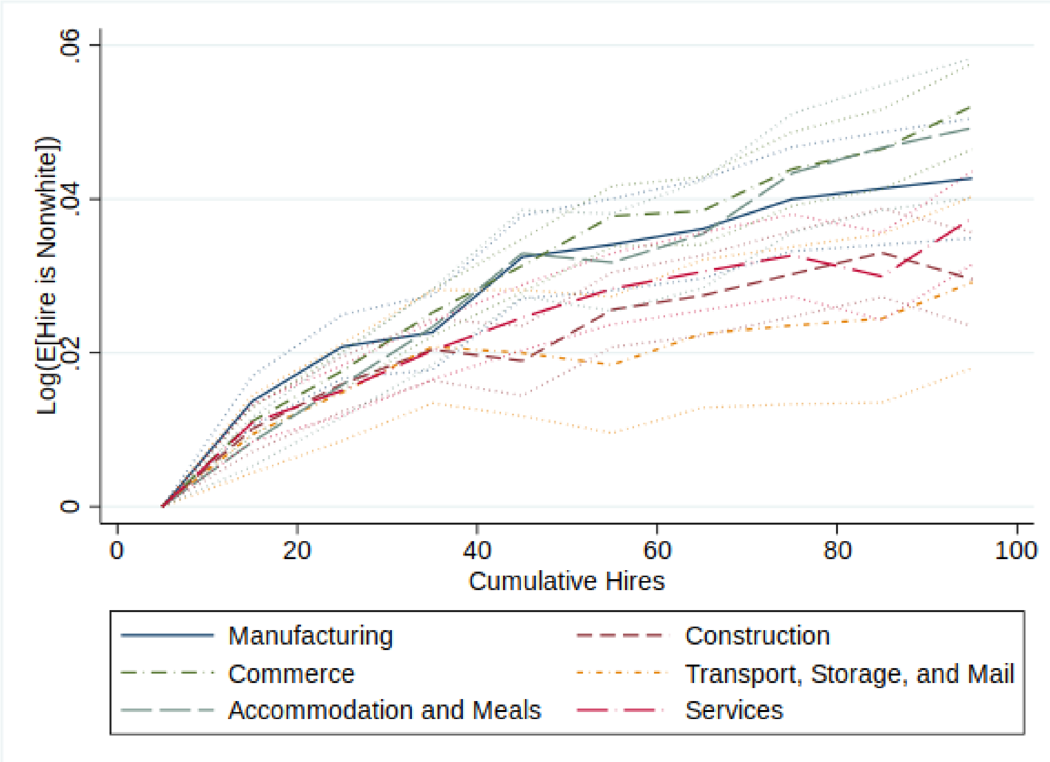
Table C.10

Racial Differences in Job Seniority, by Founder Race

	All Entrants			White Founders			Nonwhite Founders		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
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	

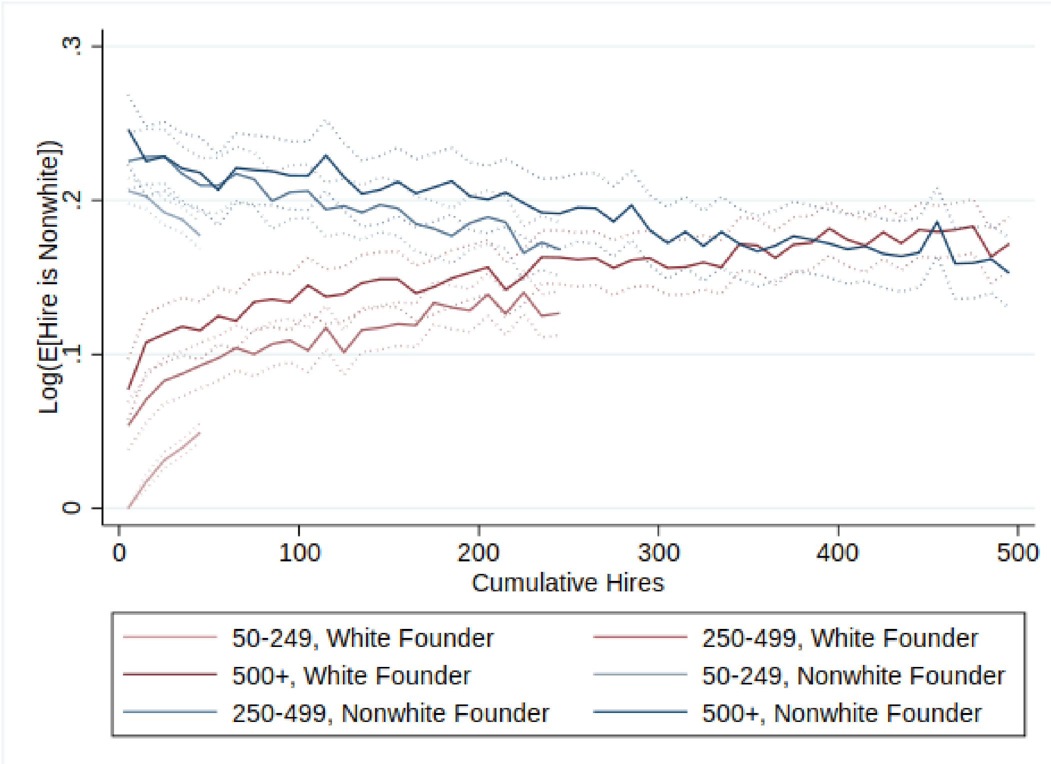
Note: This table presents regression coefficient estimates equation (C.7). The outcome is the *seniority index*, which summarizes an employee's tenure relative to their colleagues, and is defined in equation (C.6). Each observation is a job spell-year. Columns 1 through 3 pool all entrant firms, columns 4 through 6 limits to firms with white founders, and columns 7 through 9 limit to firms with nonwhite founders.

Figure C.1
Nonwhite Share of Hires by Cumulative Hires and Industry



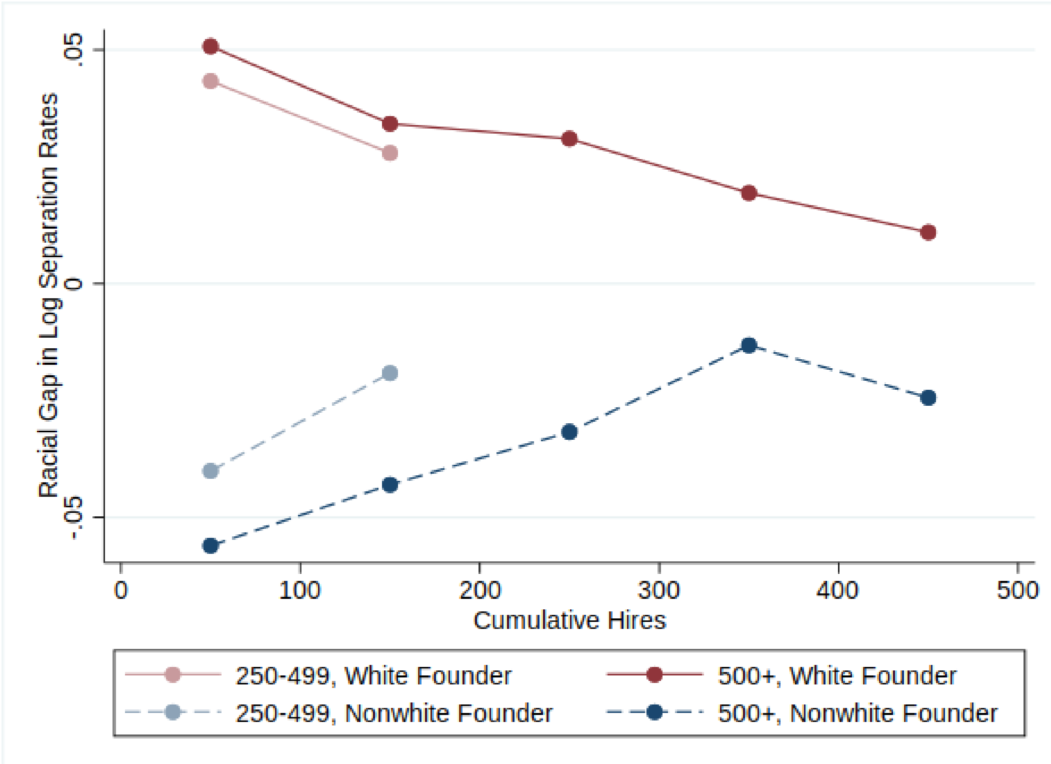
Note: This figure plots the η^n coefficient estimates from equation (??), summarizing the relationship between an establishment's racial composition of hires and its cumulative hires to date (n), separately by industry. The model is estimated via Poisson quasi maximum likelihood. We exclude the inferred founder from the new hires we consider and when measuring cumulative hires. The omitted category is the first bin of hires in each industry.

Figure C.2
Convergence in Nonwhite Share of Hires, by Ownership Race



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 $\eta^{s,n,r}$ coefficient estimates from equation (5), summarizing the relationship between a firm's racial composition of hires, its cumulative hires to date (n), and the race of its founder (r) for each firm category s. The model is estimated via Poisson quasi maximum likelihood. We exclude the inferred founder from the new hires we consider and when measuring cumulative hires. The omitted category is the first ten hires for firms with white founders and 50–249 total observed hires. Founder race is inferred from the racial composition of the firm's ownership.

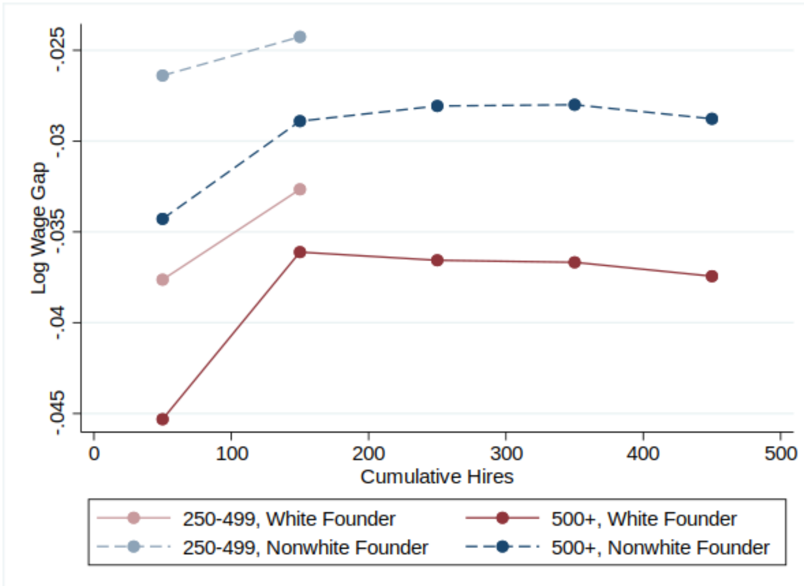
Figure C.3
Convergence in Separation Rates



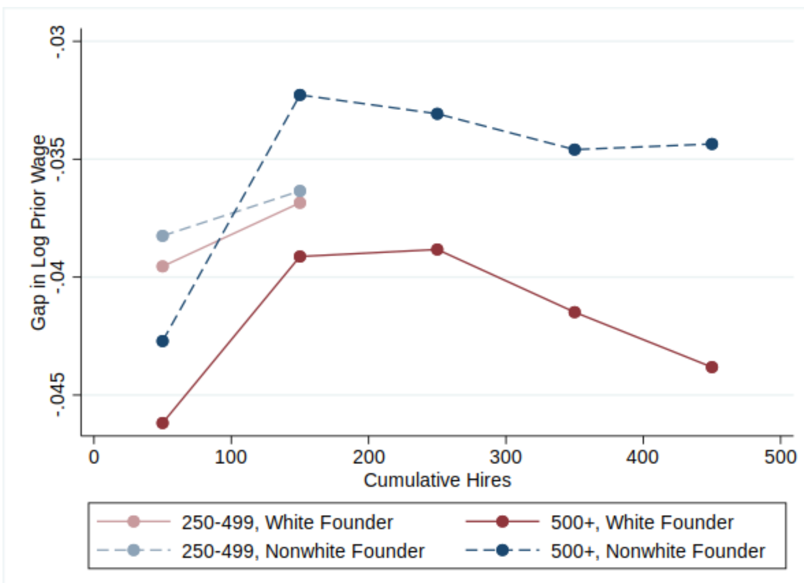
Note: This figure plots the adjusted, firm-level nonwhite-white difference in log 12-month separation rates (ψ_{jN}^{NW}) as a function of founder race and cumulative hires. Firm-specific racial differences in separation rates, which vary with cumulative hives, are constructed as described in equation (8), replacing the outcome with an indicator for separation within 12 months. The model is estimated via Poisson quasi maximum likelihood. Cumulative hires are divided into buckets of 100 hires: 1–100, 101–200, and so on, up to 401–500. The estimation sample is limited to hires 1–200 for firms with 250–499 hires and hires 1–500 for firms with at least 500 hires. Founder race is inferred from the race of the top-paid manager or employee at entry.

Figure C.4
Wages by Cumulative Hires and Founder Race

(a) Starting Log Wage



(b) Prior Starting Log Wage



Note: This figure plots the adjusted, firm-level nonwhite-white difference in log monthly wage as a function of founder race and cumulative hires. Firm-specific racial differences in dismissal rates, which can vary with cumulative hires, are constructed as described in equation (C.5). Cumulative hires are divided into buckets of 100 hires: 1–100, 101–200, and so on, up to 401–500. The estimation sample is limited to hires 1–200 for firms with 250–499 hires and hires 1–500 for firms with at least 500 hires. Founder race is inferred from the race of the top-paid manager or employee at entry.

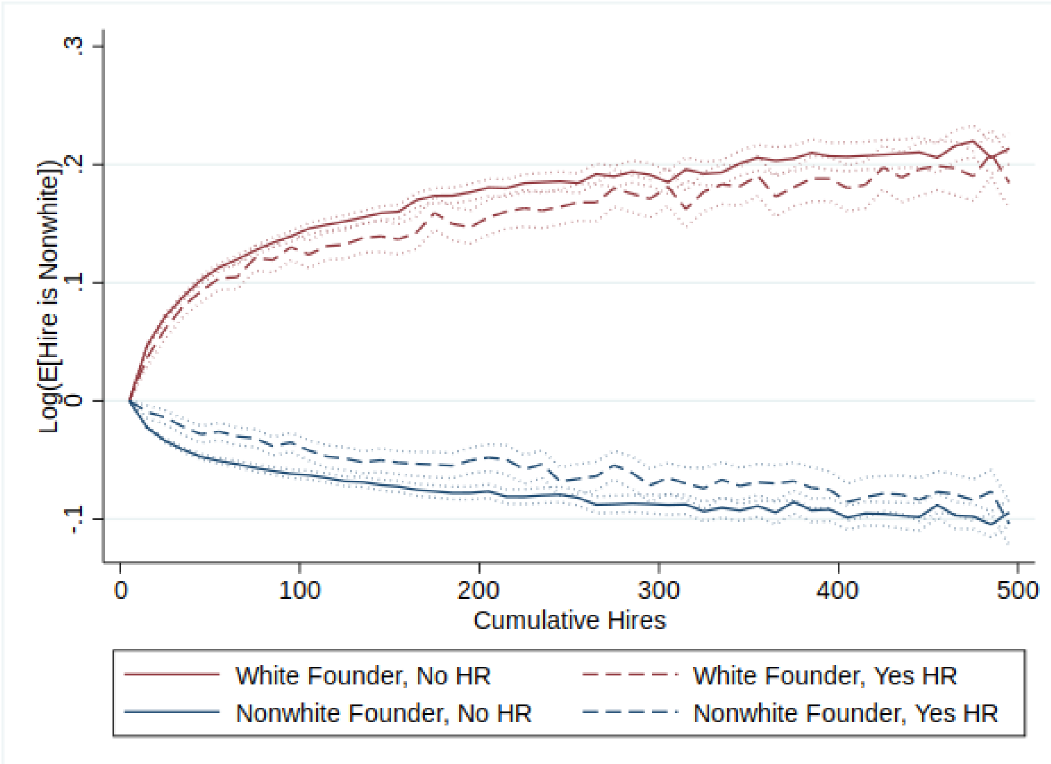
Figure C.5
Job-to-Job Mobility and Prior Quit Rates by Cumulative Hires and Race

(a) Job-to-Job Mobility

(b) Quit Prior Job

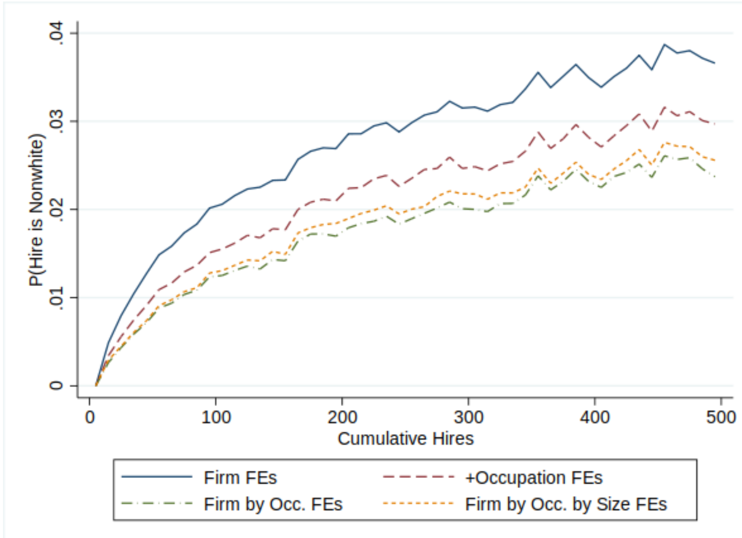
Note: This figure plots the relationship between a firm's cumulative hires to date and among new hires two measures of worker revealed preference over jobs: an indicator for whether that hire moved directly from a previous job (Panel A) and an indicator for whether that hire quit their previous job (Panel B). We plot the η coefficients from estimation of equation (12), where the model is estimated separately for pairs of worker and founder race.

Figure C.6
Nonwhite Share of Hires and Cumulative Hires by Firm HR Presence



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 (5), 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 model is estimated via Poisson quasi maximum likelihood. The omitted category is the first five hires for firms with white founders. Founder race is inferred from the race of the top-paid manager or employee at entry. Firms are categorized based on whether any observed hires are for 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).

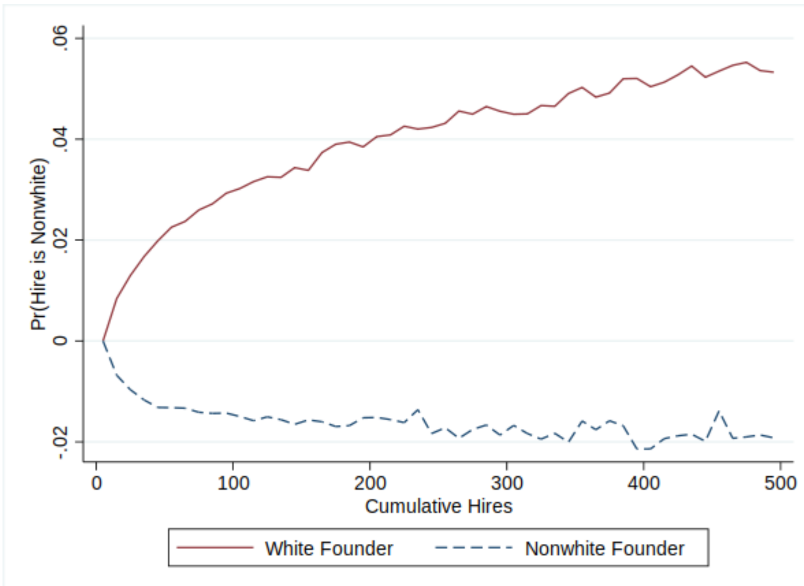
Figure C.7
Nonwhite Share of Hires Increases over Life Cycle, Linear Probability Model



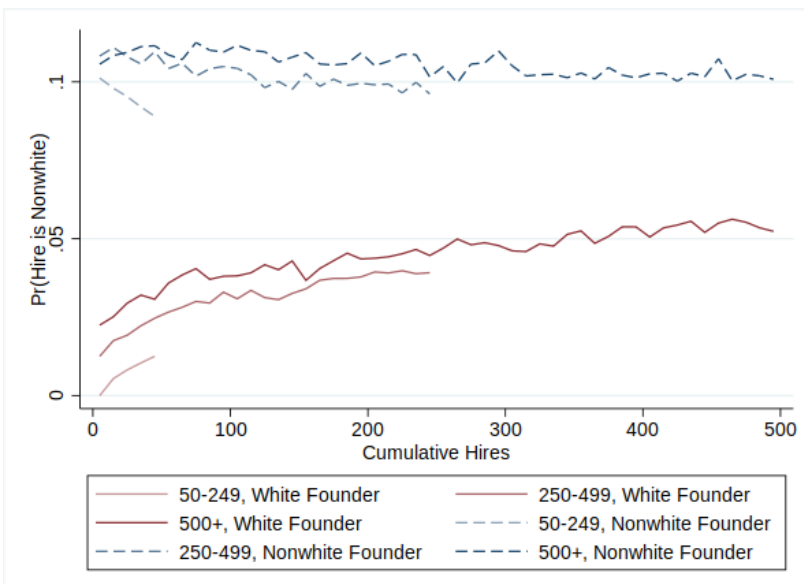
Note: This figure plots the η^n coefficient estimates from equation (C.2), summarizing the relationship between a firm's racial composition of hires and its cumulative hires to date (n). The figure includes estimates for four specifications. The baseline specification (blue) includes firm fixed effects and microregion by year fixed effects, but no additional controls. The second specification (red) replaces microregion by year fixed effects with 6-digit occupation by microregion by year fixed effects. The third (green) and fourth (orange) specifications replace firm fixed effects with firm by occupation fixed effects and firm by occupation by contemporaneous firm size fixed effects, respectively, and replaces occupation by microregion by year fixed effects with microregion by year fixed effects. We exclude the inferred founder from the new hires we consider and when measuring cumulative hires. The omitted category is the first ten hires after the year of entry.

Figure C.8
Founder Race and Convergence in Nonwhite Share of Hires, Linear Probability Model

(a) All Entrants

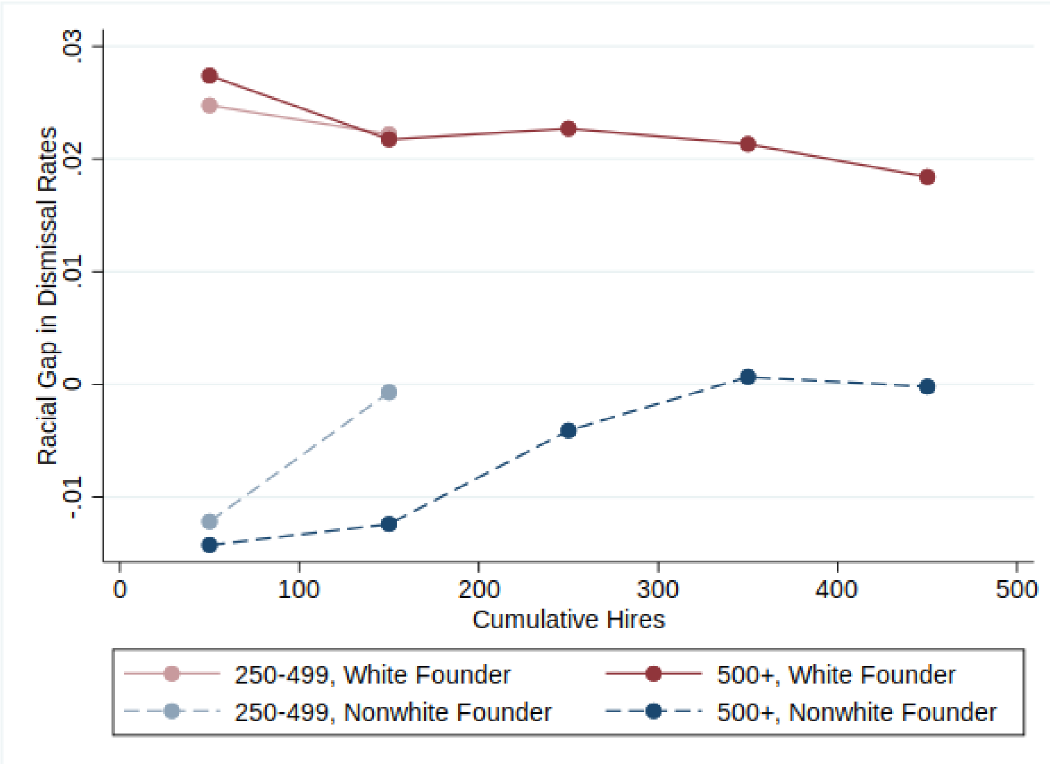


(b) Balanced Panel of Entrants



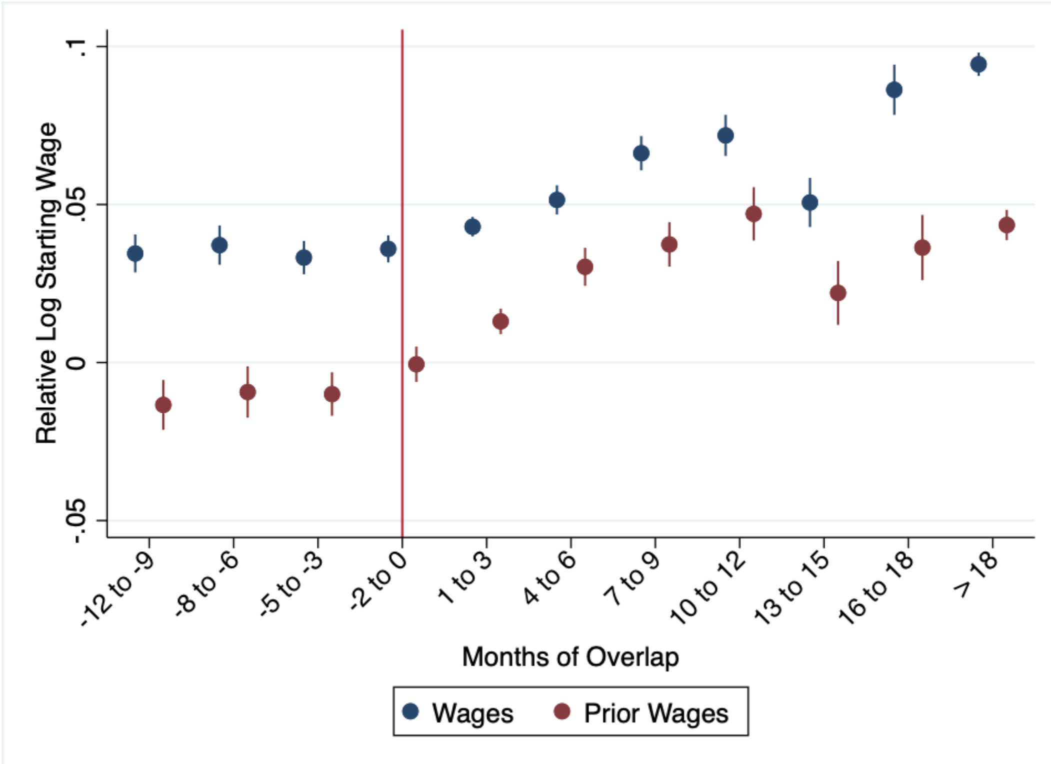
Note: This figure plots the relationship between the racial composition of a firm's hires and its cumulative hires to date. Panel A plots the $\eta^{n,r}$ coefficient estimates from equation (C.3), summarizing the relationship between a firm's racial composition of hires, its cumulative hires to date (n), and the race of its founder (r). Panel B plots the $\eta^{s,n,r}$ coefficient estimates from equation (C.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) for each firm category s. The omitted category is the ten five hires after the year of entry for establishments where the nonwhite share of the firm's incumbent employees is 0–50%. In both panels we exclude the inferred founder from the new hires we consider and when measuring cumulative hires. In Panel A the omitted category is the first ten hires. In Panel B the omitted category is the first ten hires 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 C.9
Convergence in Dismissal Rates, Linear Probability Model



Note: This figure plots the adjusted, firm-level nonwhite-white difference in 12-month dismissal rates as a function of founder race and cumulative hires. Firm-specific racial differences in dismissal rates, which vary with cumulative hires, are constructed as described in equation (C.4). The model is estimated via Poisson quasi maximum likelihood. Cumulative hires are divided into buckets of 100 hires: 1–100, 101–200, and so on, up to 401–500. The estimation sample is limited to hires 1–200 for firms with 250–499 hires and hires 1–500 for firms with at least 500 hires. Founder race is inferred from the race of the top-paid manager or employee at entry.

Figure C.10
Connected Hires Are Paid More



Note: This figure plots the θ^m coefficient estimates as described in equation (11), replacing the outcome with log starting wage. The θ^m coefficients identify the log starting wages of placebo connected and connected hires as a function of overlap relative to hires who are neither.