Co-Racial Hiring: Firm Dynamics and Aggregate Consequences

Conrad Miller*

Ian Schmutte†

May 2026

Abstract

A firm's incumbent workforce can tilt subsequent hiring toward racial groups already overrepresented among its employees, a phenomenon we refer to as co-racial hiring. We study its implications for firms' hiring over the life cycle and aggregate racial differences in employment. Co-racial hiring could, in principle, reproduce a firm's existing racial composition and entrench workplace segregation. We show in a simple model that self-perpetuating segregation is a limiting case rather than an inevitable consequence: when recruiting reflects both incumbent composition and the broader pool of external job seekers, initial composition differences can fade as firms accumulate hires. We test the model's predictions using linked employer-employee data from Brazil, where nonwhite workers are underrepresented in formal employment and among entrepreneurs. For the same job, later hires are more likely to be nonwhite than earlier hires. At entry, firms with white founders are about 30% less likely to hire nonwhite workers than comparable firms with nonwhite founders, but this gap shrinks with cumulative hires and is nearly eliminated after 400 hires. Racial differences in dismissal rates follow a similar pattern. We then embed co-racial hiring in an equilibrium search model with group-specific matching. Interpreted through the model, the firm life-cycle pattern we document—that the nonwhite share of hires rises with cumulative hires—implies that co-racial hiring raises nonwhite unemployment relative to white unemployment.

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

†Amazon and 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

A growing literature shows that the racial composition of a firm's incumbent workforce can affect the composition of its subsequent hires. We refer to this phenomenon as *co-racial hiring*. This pattern may arise because employees tend to refer candidates from their own group (Fernandez et al., 2000; Petersen et al., 2000; Fernandez and Sosa, 2005; Dustmann et al., 2016), because hiring managers may prefer or better screen workers from their own group (Stoll et al., 2004; Giuliano et al., 2009; Åslund et al., 2014; Benson et al., 2022), or because workers may be more productive under co-racial supervision (Hjort, 2014; Glover et al., 2017). This literature has largely focused on the short-term, within-firm implications of these mechanisms—for example, how the identity of the manager or referring employee affects the identity of the next hire. We take this evidence as our starting point and ask how co-racial hiring shapes firms' workforce composition over the life cycle and racial differences in employment in the aggregate.

The broader consequences of these firm-level tendencies are not obvious. Co-racial hiring could, in principle, reproduce a firm's existing racial composition and entrench workplace segregation. In that case, founder race would have lasting effects because it shapes the firm's initial social conditions. But the founder's influence need not be permanent: when hiring remains partly anchored to the broader pool of job seekers, founder-linked differences can fade over time. The aggregate implications are also ambiguous. Co-racial hiring may amplify racial inequality, but there are at least two reasons it need not. First, its effects could offset across firms: different firms tilt hiring toward different groups. Second, labor markets adjust. If one group is under-targeted by firms, workers from that group become more available to hire.

We formalize co-racial hiring in a model in which firms' recruiting depends on both incumbent composition and the composition of the external job seekers. The model highlights two cases. If recruiting simply mirrors incumbent composition, initial racial differences across firms are reproduced indefinitely. If recruiting remains partly anchored to the external market, those differences fade as firms accumulate hires. The structure is deliberately agnostic about the mechanism generating co-racial hiring.

In the latter case, where recruiting remains partly anchored to the external market, the model yields four empirical predictions. First, founder-linked initial conditions matter most early in the firm life cycle. Second, otherwise comparable firms with founders from different racial groups converge in the racial composition of their hires as cumulative hires increase. The third prediction generalizes this convergence logic: dispersion across firms in the expected racial composition of hires declines over the life cycle. Fourth, incumbent composition is the relevant state variable, so conditioning on incumbent composition attenuates the relationship between founder type, cumulative hires, and the race of the next hire. Similar dynamics arise in dismissals if co-racial hires are less likely to be poor matches ex post or less likely to be marginal hires (Benson et al., 2022).

We test these predictions using linked employer-employee data from Brazil, a setting where nonwhite workers are underrepresented in formal employment and among entrepreneurs (IBGE, 2021).

We first show that, within the same job, later hires are more likely to be nonwhite than earlier hires. This pattern is not explained by changes in observable job characteristics, including detailed occupation and contemporaneous firm size. We then show that this life-cycle profile differs sharply by founder race, consistent with the model. We proxy founder race by the race of the firm's top-paid manager or owner. At entry, firms with white founders are substantially less likely to hire nonwhite workers than otherwise similar firms with nonwhite founders. Within the same local labor market and occupation, early hires at white-founded firms are about 30% less likely to be nonwhite. This gap shrinks with cumulative hires: after 400 hires, the next hire at a white-founded firm is only about 4% less likely to be nonwhite.

We find similar patterns for new establishments opened by preexisting firms. We proxy for initial conditions using the racial composition of the parent firm's incumbent workforce. New establishments opened by firms with a higher white share among incumbent employees are initially more likely to hire white workers, but these differences also shrink with cumulative hires. More generally, dispersion across firms in the racial composition of hires declines over the life cycle. Consistent with the model's mechanism, conditioning on the racial composition of incumbent employees substantially attenuates the relationship between cumulative hires, founder race, and the race of the next hire.

Dismissals follow analogous dynamics. At firms with white founders, nonwhite employees are dismissed at higher rates, while the reverse is true at firms with nonwhite founders. These racial gaps in dismissal rates also shrink as cumulative hires increase, as does dispersion across firms in dismissal gaps.

Although the influence of founder race diminishes with cumulative hires, few firms make enough hires to reach their steady-state composition. We estimate that, for the average entrant firm in our sample, the nonwhite share of hires is 4% below the plateau value. This suggests that transition dynamics matter in the aggregate: if most firms remain small or young, founder-linked initial conditions continue to shape hiring over much of the life cycle, so racial differences in entrepreneurship can generate racial inequality in labor demand.

While suggestive, the firm life-cycle profiles do not mechanically translate into aggregate unemployment differences. Labor markets also adjust: under-targeted groups become more represented among job seekers, and vacancies aimed at them become easier to fill. These forces can raise the group's share of completed hires at any given recruiting allocation, even while firms with different incumbent compositions continue to hire differently. To capture these forces, we embed co-racial hiring in a search model in which operating firms post a fixed number of vacancies per period and choose how to direct those vacancies across white and nonwhite job seekers. Founder type matters through the firm's initial recruiting orientation: early vacancies at white-founded and nonwhite-founded firms are targeted differently. Completed hiring also depends on group-specific job-filling rates, so targeted recruiting and completed hires can differ when one group is easier to hire.

The model delivers a simple threshold result. Co-racial hiring raises nonwhite unemployment relative to white unemployment when nonwhite founders are underrepresented relative to the nonwhite share of the labor force. The reason is that market-clearing forces are stabilizing but self-limiting. When nonwhite workers are under-targeted, these forces push firms back toward nonwhite hiring. But the response weakens as unemployment rates converge. At parity, the unemployed pool mirrors the labor force and vacancy-filling rates are equal across groups, so there is no remaining market-clearing adjustment pushing firms toward nonwhite workers. If nonwhite founders are underrepresented at that point, the founder imbalance still leads firms to direct too few vacancies toward nonwhite job seekers, so parity cannot be an equilibrium.

The rising life-cycle profile links this threshold result to the data. Because the nonwhite share of hires rises with cumulative hires, the average entrant firm begins below its steady-state nonwhite hire share. In the model, this occurs exactly when nonwhite founders are underrepresented relative to the nonwhite labor force share. Hence racial differences in entrepreneurship can generate persistent racial differences in labor demand and unemployment, even though founder effects fade within individual firms.

Publicly available U.S. data provide suggestive evidence of a similar increasing life-cycle profile for Black hires. We use Census tabulations to construct aggregate firm cohorts defined by implied entry year, geography, and industry. In these aggregate cohorts, the Black share of new hires rises as firms age. These tabulations do not identify cumulative hires within firms, so the evidence is necessarily indirect; nevertheless, the pattern is qualitatively consistent with the life-cycle profile we document in Brazil.[1]

A substantial empirical literature documents patterns consistent with co-racial (or co-ethnic) hiring. One set of papers studies referral networks and shows that incumbent workers influence who is recruited and hired in ways that correlate with the firm's existing social composition (Fernandez et al., 2000; Petersen et al., 2000; Fernandez and Sosa, 2005; Dustmann et al., 2016). A second set of papers studies managers and shows that hiring and personnel outcomes depend on the demographic match between managers and workers (Giuliano et al., 2009, 2011; Benson et al., 2022; Bossler et al., 2020; Hensvik, 2014; Cullen and Perez-Truglia, 2023). A third set of papers documents that entrepreneurs and business owners disproportionately employ workers from their own racial or ethnic group (Bates, 2006; Boston, 2006; Dias and Rocha, 2021; Kerr and Kerr, 2021). Across these settings, firms draw hires in a manner that depends on the composition of the incumbent workforce.

We take that regularity as a starting point and ask what it implies for firm dynamics and aggregate inequality. The main contribution is not to establish that co-racial hiring exists, but to consider how founder-linked initial conditions propagate through subsequent hiring, how quickly those effects decay, and how those within-firm dynamics map into aggregate racial differences in hiring and unemployment.

This paper is broadly related to prior work that considers the dynamic consequences of in-group preferences in social interactions. Co-racial hiring can, in principle, lead to tipping or extreme

[1] This pattern is also consistent with earlier U.S. evidence that the Black share of employment is increasing in employer size, except among Black-run businesses (Holzer, 1998; Miller, 2017), where the pattern is reversed.

segregation (Becker, 1957; Schelling, 1971; Pan, 2015), but it need not. Referral-based hiring, for example, need not produce permanent segregation if referral networks are themselves imperfectly segregated, or if workers who were initially hired through market channels later become sources of referrals (Rubineau and Fernandez, 2013, 2015). Whereas this literature focuses on the dynamics of specific mechanisms, we take a more "reduced form" and mechanism-agnostic approach and study what these dynamics imply for within-firm convergence and aggregate inequality.

Finally, we contribute to the literature on racial differences in unemployment. Search models with employer prejudice or statistical discrimination can generate group differences in job-finding and unemployment even when wage differences are modest (Black, 1995; Bowlus and Eckstein, 2002; Rosén, 2003; Lang and Lehmann, 2012). The U.S. evidence poses a particular puzzle because, relative to the wage gap, the Black-white unemployment gap is not well explained (in a statistical sense) by standard premarket skill measures like test scores (Ritter and Taylor, 2011). Our model highlights a distinct mechanism: when groups differ in entrepreneurship rates, founder-linked hiring dynamics within firms can translate those differences into persistent inequality in hiring and unemployment, even when the founder effect within a given firm fades over time.

Section 2 develops the model of within-firm co-racial hiring and its implications for founder effects and workforce dynamics. Section 3 describes the Brazilian context and employer-employee data that form the basis of our study. Section 4 presents evidence on hiring and dismissals over the firm life cycle. Section 5 embeds the within-firm mechanism in a search model, characterizes the stationary equilibrium, and discusses the model's stylized implications for racial inequality in formal sector employment. Section 6 concludes.

2 A model of co-racial hiring within the firm

We begin with a within-firm model of hiring dynamics. The goal is to isolate how incumbent composition propagates through subsequent hiring while holding fixed the composition of external job seekers. The aggregate model in section 5 builds on the same structure and endogenizes the composition of the external market.

Workers belong to one of two groups, $g \in \{W, N\}$, where N denotes nonwhite and W denotes white. Let $\gamma \in [0, 1]$ denote the nonwhite share among the relevant pool of external job seekers. In this section, γ is exogenous.

A firm is indexed by founder type $\tau \in \{W, N\}$. Time is indexed by completed hires, $k = 0, 1, 2, \ldots$. Let $\pi_k \in [0, 1]$ denote the firm's composition state after k hires. When $k \geq 1$, π_k is the nonwhite share of incumbent employees. When $k = 0$, π_0^τ is a founder type-specific orientation rather than a literal workforce share. We set $\pi_0^W = 0$ and $\pi_0^N = 1$. Entrant firms begin with no employees but inherit a founder-linked orientation toward one group or the other.[2]

Let $Y_{k+1} \in \{0, 1\}$ indicate whether completed hire $k+1$ is nonwhite. We distinguish between two margins in the hiring process. The first is *recruiting effort*, which the firm allocates between groups.

[2]None of the qualitative results below depends on this exact choice; in Appendix B, we state results for arbitrary $\pi_0^W < \pi_0^N$.

We use this term broadly to mean actions that shape which candidates are reached, considered, or prioritized. The second margin is *vacancy filling*: effort directed toward a group converts into a completed hire at a group-specific rate. Completed hiring therefore depends both on how the firm allocates recruiting effort and on how easily each group's targeted recruiting converts into hires.

Appendix B.2 argues that several co-racial hiring mechanisms—including referral hiring, screening discrimination, taste-based discrimination, and production complementarities—can be represented by making the share of recruiting effort directed toward nonwhite job seekers affine in the firm's incumbent composition:

$$p(\pi_k, \gamma) = a(\gamma) + b(\gamma)\pi_k.$$

For the main analysis, we focus on the following benchmark linear case:

$$p(\pi_k, \gamma) = (1-\rho)\gamma + \rho\pi_k, \qquad \rho \in [0, 1].$$

The parameter ρ measures the strength of co-racial hiring, or equivalently the weight placed on incumbent composition. In the limiting cases where $\rho = 0$ or $\rho = 1$, recruiting mirrors the external pool or incumbent composition, respectively. We focus on the interior case where $0 < \rho < 1$ and recruiting reflects both incumbent composition and the external pool.

Referral hiring provides a direct microfoundation for this benchmark. Suppose a fraction ω of recruiting is mediated by incumbent referrals and a fraction $1 - \omega$ comes through the external market. Suppose further that referral networks are partly segregated, so the nonwhite share of referral candidates generated by incumbents is

$$(1-\eta)\gamma + \eta\pi_k, \qquad \eta \in [0, 1].$$

Then the nonwhite share of recruiting is

$$p(\pi_k, \gamma) = \omega\big[(1-\eta)\gamma + \eta\pi_k\big] + (1-\omega)\gamma = (1-\rho)\gamma + \rho\pi_k, \qquad \rho \equiv \omega\eta.$$

The edge case $\rho = 1$ corresponds to $\omega = \eta = 1$: all recruiting is mediated by incumbent referrals and referral networks are fully segregated.

We focus on the benchmark linear case in the main text because its transition dynamics are especially transparent. Appendix B.2 explains how the results extend to the broader affine class.

Targeted recruiting need not coincide with completed hiring. The firm faces constant group-specific filling rates $q_N > 0$ and $q_W > 0$. These capture how easily recruiting aimed at nonwhite and white job seekers converts into completed hires. If the firm directs share $p(\pi_k, \gamma)$ toward nonwhite job seekers and share $1 - p(\pi_k, \gamma)$ toward white job seekers, then the probability that the next completed hire is nonwhite is

$$\Pr(Y_{k+1} = 1 \mid \pi_k, \gamma, q_N, q_W) = \widehat{p}(\pi_k, \gamma; q_N, q_W) \equiv \frac{p(\pi_k, \gamma)q_N}{p(\pi_k, \gamma)q_N + [1 - p(\pi_k, \gamma)]q_W}.$$

The distinction between $p(\pi_k, \gamma)$ and $\widehat{p}(\pi_k, \gamma; q_N, q_W)$ becomes more important in the aggregate model, where q_N and q_W are endogenous. The first object captures the firm's recruiting allocation; the second captures the composition of completed hires. If one group is easier to hire from, completed hiring is tilted toward that group for a fixed recruiting allocation.

A useful special case is equal filling rates, $q_N = q_W$. Then completed hiring coincides with recruiting:

$$\widehat{p}(\pi_k, \gamma; q_N, q_W) = p(\pi_k, \gamma).$$

The firm's composition evolves according to

$$\pi_{k+1} = \frac{k\pi_k + Y_{k+1}}{k+1}.$$

Hence conditional mean composition satisfies

$$\mathbb{E}[\pi_{k+1} \mid \pi_k, \gamma, q_N, q_W] = \pi_k + \frac{\widehat{p}(\pi_k, \gamma; q_N, q_W) - \pi_k}{k+1}. \tag{1}$$

Current composition matters because it shapes how recruiting is allocated, and filling rates determine how that allocation translates into actual hires.

We impose the following contraction condition:

$$\kappa \equiv \rho \, \frac{\max\{q_N, q_W\}}{\min\{q_N, q_W\}} < 1. \tag{2}$$

Under (2), the map $\widehat{p}(\cdot, \gamma; q_N, q_W)$ is a contraction on $[0, 1]$. It therefore has a unique fixed point $\pi^*(\gamma, q_N, q_W) \in [0, 1]$ satisfying

$$\pi^*(\gamma, q_N, q_W) = \widehat{p}(\pi^*(\gamma, q_N, q_W), \gamma; q_N, q_W).$$

Founder type affects the path through the initial condition, but not the long-run target. Firms with different founders begin from different orientations, but their hiring dynamics drift toward a common target determined by the fixed environment (γ, q_N, q_W).

When $q_N = q_W$, the contraction condition is automatic because $\kappa = \rho < 1$. Completed hiring equals recruiting, the fixed point is $\pi^*(\gamma, q_N, q_W) = \gamma$, and (1) becomes

$$\mathbb{E}[\pi_{k+1} \mid \pi_k, \gamma] = \pi_k + \frac{(1-\rho)(\gamma - \pi_k)}{k+1}.$$

For any $\rho < 1$, founder effects are therefore strongest early in the firm's life cycle and decay as cumulative hires increase and incumbent composition moves toward the external pool.

2.1 Testable implications

The model yields four central implications for hiring.

1. **Founder-group early hiring.** Early hires are disproportionately drawn from the founder's group.

2. **Convergence in cumulative hires.** Differences in hiring between firms with white and nonwhite founders shrink with cumulative hires.

3. **Declining between-firm dispersion.** For firms facing a common external composition γ and the same filling rates (q_N, q_W), hire composition converges toward a common target $\pi^*(\gamma, q_N, q_W)$. Therefore, between-firm dispersion in conditional mean hire composition declines with cumulative hires.

4. **Incumbent composition as the state variable.** Conditioning on incumbent composition attenuates the relationship between founder type, cumulative hires, and the racial composition of new hires.

We test each of these predictions in section 4.

Although the model above abstracts from employee turnover, the same microfoundations also imply auxiliary predictions for turnover (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—as would be the case for screening discrimination and common formulations of referral hiring (Montgomery, 1991; Simon and Warner, 1992; Brown et al., 2016; Pallais and Sands, 2016; Burks et al., 2015; Dustmann et al., 2016). With taste-based discrimination or production complementarities, co-racial hires are less likely to be marginal. In either case, co-racial hires are less likely to be dismissed.[3] Early in the firm's life cycle, founder race proxies for the composition of incumbents, so this logic implies three additional predictions.

5. **Founder-linked dismissal gaps.** At white-founded firms, recent nonwhite hires are more likely to be dismissed than recent white hires; at nonwhite-founded firms, the opposite holds.

6. **Convergence in dismissal gaps.** Differences in racial dismissal gaps between firms with white and nonwhite founders shrink with cumulative hires.

7. **Declining between-firm dispersion in dismissal gaps.** Between-firm dispersion in racial dismissal gaps declines with cumulative hires.

3 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., 2021). However, Brazil has few regulations that protect workers against employment discrimination in the private

[3]See Benson et al. (2022) for more details on these predictions.

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.

3.1 Legal and social context

Brazil was founded as a race-based slave society and has persistent racial disparities across many socio-economic outcomes. For many decades after the end of slavery, Brazil maintained a national myth that it was a "racial democracy" in which racial disparities were incidental and transitory (Fiola, 1990). Brazil did not construct explicitly racist legal institutions equivalent to the Jim Crow era in the U.S., did not prohibit racial intermarriage, and did not 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. (2021), 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 Pesquisa Nacional por Amostra de Domicílios (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. The share unemployed is 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,

[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. (2021).

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 Domicílios (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.

while 2.1% and 1.8% of brown and black men are entrepreneurs.

Among private sector employees, whites have more years of 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 somewhat lower for nonwhite workers.

3.2 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 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 3.3 below for a description of the ownership data). 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.

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 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.

3.3 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 Brasil.

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.

4 Co-racial hiring evidence

We begin by documenting a robust life-cycle pattern in hiring composition: within entrant firms, later hires are more likely to be nonwhite than earlier hires. This pooled pattern is not itself a sharp prediction of the model, because the model does not sign the average slope across firms. But

[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

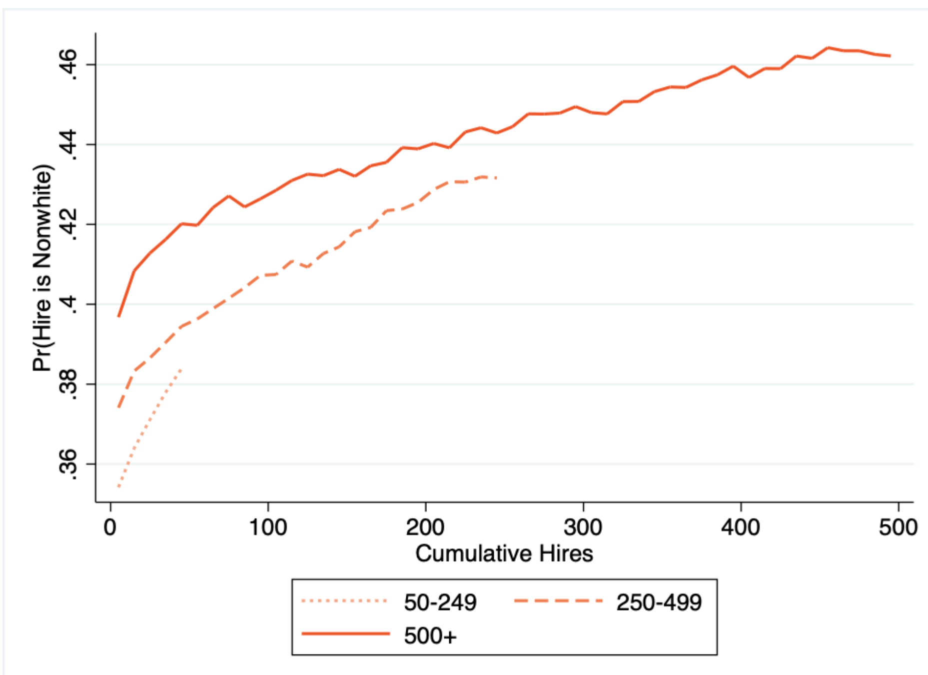
This table reports summary statistics for entrant HQ establishments in the *Relação Anual de Informações Sociais* (RAIS) data for the years 2003–2017. In columns 1 through 3 we infer the race of the firm's founder using the race of the top-paid manager (or top-paid employee if there is no manager present) 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.

it is consistent with the model if entrant firms begin below their target nonwhite share on average. We then turn to the model's sharper hiring predictions in section 4.2: founder-linked differences in hiring are largest early in the firm life cycle and shrink with cumulative hires, conditioning on incumbent composition attenuates life-cycle hiring patterns, and dispersion across firms declines as cumulative hires increase. Section 4.3 presents auxiliary evidence on dismissals.

4.1 A motivating stylized fact

We begin by documenting a 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.[10]

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

One potential explanation for this pattern is that the types of job vacancies firms fill change

[10]This statistic refers to 18–65 year olds with at least one year of potential work experience and is based on Pesquisa Nacional por Amostra de Domicílios (PNAD) household survey data from 2003 to 2015. See section 3.1 for more details.

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. 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

$$\log(E(\text{NONWHITE}_{jh}|\cdot)) = \sum_n \eta^n \times \mathbb{1}_{\{N(j,h)=n\}} + \tau_{t(j,h)} + \psi_j + X_{jh} \tag{3}$$

via Poisson quasi maximum likelihood (Correia et al., 2020), where each observation is a new hire.[11] NONWHITE_{jh} is an indicator for whether hire h at firm j is nonwhite. $N(j,h)$ groups firm j's hires into bins.[12] 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.[13] 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

[11] Note that the fixed effects Poisson estimator only invokes the conditional mean assumption in (3) 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.

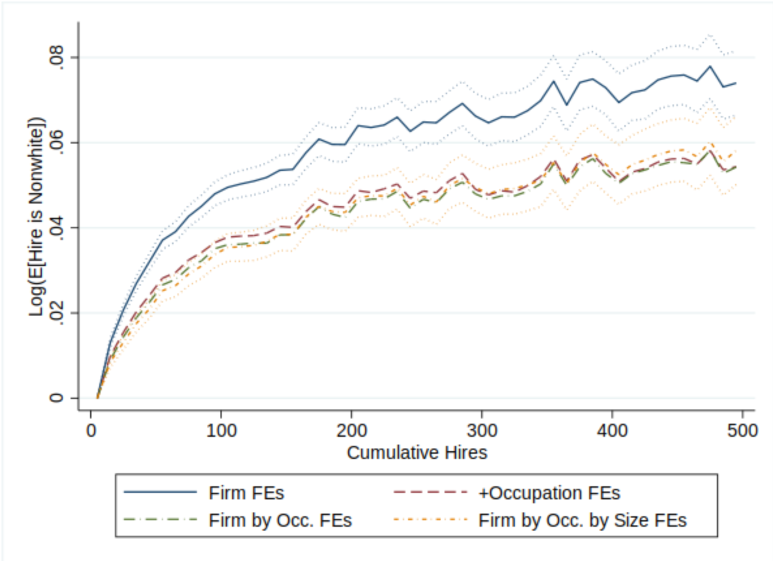
[12] For new hires with the same start date, ties are broken randomly.

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

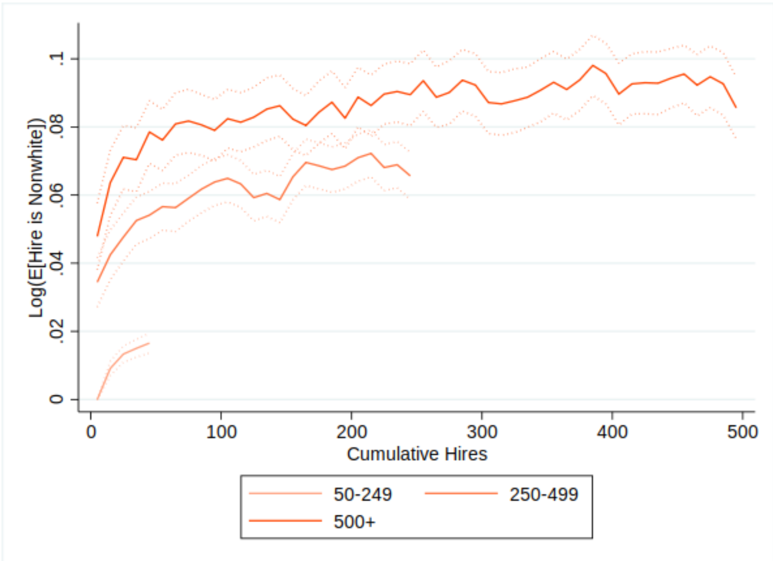
Figure 2

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 (3), 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 pointwise 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 (4), 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 pointwise 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 observed hires. In Panel B the omitted category is the first ten observed hires for firms with 50–249 total observed hires.

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 (3) for a balanced panel of firms. We also allow the η coefficients to vary with a firm's total observed hires. Specifically, we estimate

$$\log(E(\text{NONWHITE}_{jh}|\cdot)) = \sum_s \sum_n \eta^{s,n} \times \mathbb{1}_{\{S(j)=s\}} \times \mathbb{1}_{\{N(j,h)=n\}} + \tau_{t(j,h)} + \mu_{m(j)} + \omega_{o(j,h)}, \quad (4)$$

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, 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.[14] 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.[15]

In summary, the positive relationship between a firm's cumulative hires to date and the nonwhite share of its hires is a robust descriptive feature of the data. Observable job characteristics, including detailed occupation and firm size, explain only a small share of it.

We use the η^n coefficients depicted in Panel A of Figure 2 to calculate the expected deviation between a firm's nonwhite share of hires and its plateau nonwhite share, 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, let H_j denote the number of observed hires at firm j. As in equation (3), cumulative hires are grouped into bins of size 10. Let $N(j,h) \in \{10, 20, 30, \ldots\}$ denote the endpoint of the bin containing hire h. Hence $N(j,h) = 10$ for hires 1–10, $N(j,h) = 20$ for hires 11–20, and so on. Recall that we normalize the first coefficient $\hat{\eta}^{10}$ to zero.

Let $\bar{\eta}$ denote the plateau value. In this calculation, we take $\bar{\eta}$ to be the coefficient for the bin containing the 400th hire: $\bar{\eta} = \hat{\eta}^{400}$. For hires in later bins, we assign the plateau value, so that

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

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

$\hat{\eta}^N = \bar{\eta}$ for all bin endpoints $N \geq 400$. The deviation for firm j is

$$\Delta_j = \bar{\eta} - \frac{1}{H_j} \sum_{h=1}^{H_j} \hat{\eta}^{N(j,h)}.$$

Δ_j compares the plateau coefficient to the average coefficient for the hires observed at firm j. Averaging Δ_j across entrant firms gives the reported aggregate deviation.

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

4.2 Hiring evidence

We next turn to sharper tests of the model's predictions for transition dynamics. A key prediction of the model is that firms with white and nonwhite founders initially differ in the racial composition of their hires, but that these differences shrink over the firm life cycle.

To test this prediction, we estimate the following variant of (4), where we allow the $\eta^{s,n}$ coefficients to vary with the race of the firm's founder:

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

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 3. 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.1; the results are similar.) The pattern matches the model's prediction. Early in the firm life cycle, the racial composition of new hires is strongly related 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 than at firms with a white founder. The gap declines steeply with cumulative hires: by the 50^{th} hire it falls to about 15 log points, and by the 200^{th} hire to about 10 log points. After the 400^{th} hire, the gap hovers between 3 and 5 log points. Observably similar firms with white and nonwhite founders therefore converge substantially over the firm life cycle, even though small differences remain late in the life cycle.[16]

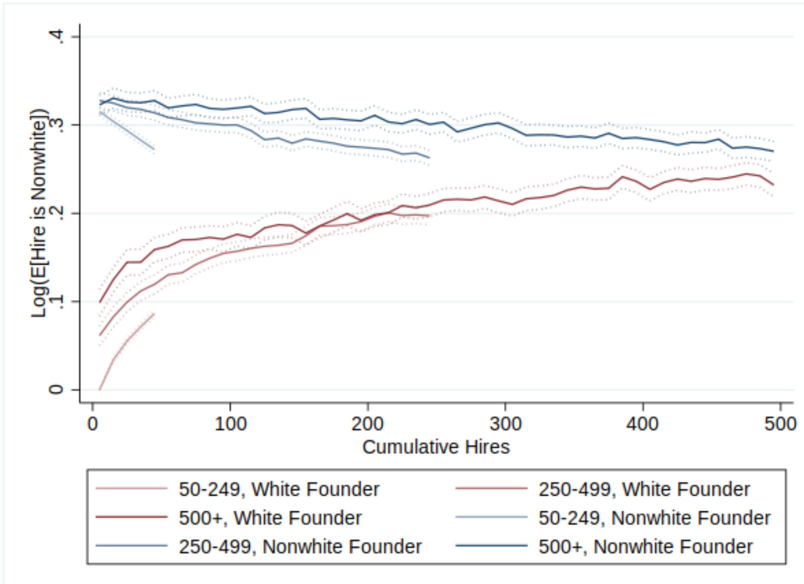
One feature of the descriptive results remains outside the scope of the co-racial hiring model in section 2. Holding cumulative hires fixed, the probability that a hire is nonwhite is increasing in the total number of hires the firm will make. This is true for both firms with white founders and firms with nonwhite founders, suggesting that co-racial hiring cannot by itself explain the pattern. We leave further exploration of this fact to future research.

[16] As in the descriptive exercises above, the patterns depicted in Figure 3 are insensitive to the job characteristics we condition on (see Appendix C and Appendix Table C.6).

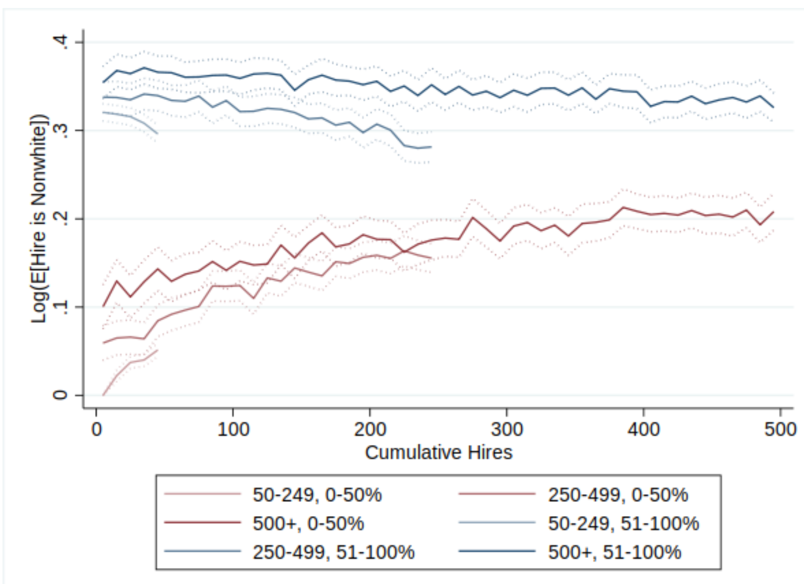
Figure 3

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%.

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 3. 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 next test a further implication of the model: conditioning on π_{jh}, the nonwhite share of incumbent employees at the firm, should substantially attenuate the relationship between cumulative hires and the nonwhite share of hires. In the model, incumbent composition is the state variable that mediates the life-cycle transition, so conditioning on π_{jh} removes most of the variation generated by cumulative hires. We estimate regression models of the form

$$\log(E(\text{NONWHITE}_{jh}|\cdot)) = \delta \log N(j,h) + \tau_{t(j,h)} + \mu_{m(j)} + \omega_{o(j,h)} + \sigma_{S(j)} + f(\pi_{jh}) \tag{6}$$

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.

We focus on convergence between firms with white and nonwhite founders, and between new establishments whose parent firms differ in incumbent composition, but the same transition logic has a broader implication. As firms accumulate hires and incumbent composition converges, dispersion across firms in the expected racial composition of hires should also decline.

We extend (5) and estimate the following model:

$$\log(E(\text{NONWHITE}_{jh}|\cdot)) = \tau_{t(j,h)} + \omega_{o(j,h)} + \theta_{jN(j,h)} \tag{7}$$

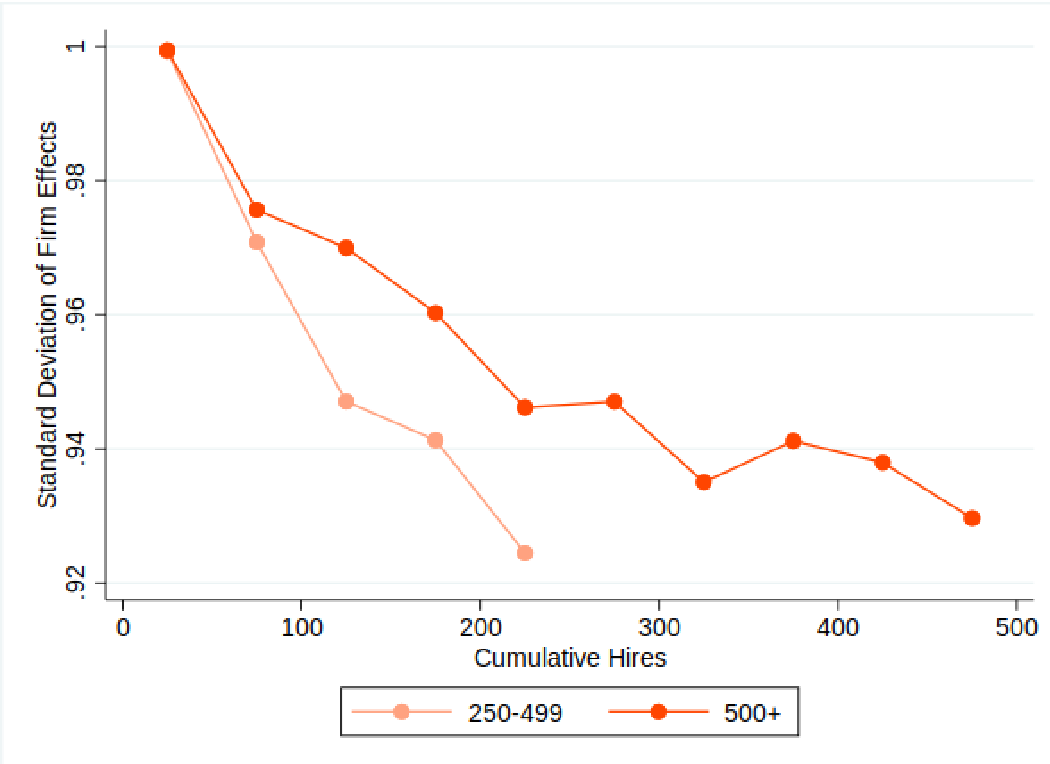
[17] For descriptive statistics on new subsidiary establishments and their hires, see Appendix Table C.4 and Appendix Table C.5.

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 (6), 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 4
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 4.2. 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).

where $\theta_{jN(j,h)}$ 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 4 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).

4.3 Dismissals evidence

We next present auxiliary evidence on dismissals. The firm-level model in section 2 is a model of hiring rather than separations, so we view the results in this subsection as complementary evidence on related forces rather than as direct tests of the model. 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

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

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 5 depicts the averages of $\psi^{NW}_{jN(j,h)}$ by cumulative hire bin separately for firms with white and nonwhite founders. For the first 100 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 401–500. By contrast, for the first 100 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 200 hires.[19]

More generally, the dispersion in firm-specific dismissal rate differences is decreasing in cumulative hires. Figure 6 plots the standard deviation of $\psi^{NW}_{jN(j,h)}$ 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).

4.4 Worker preferences

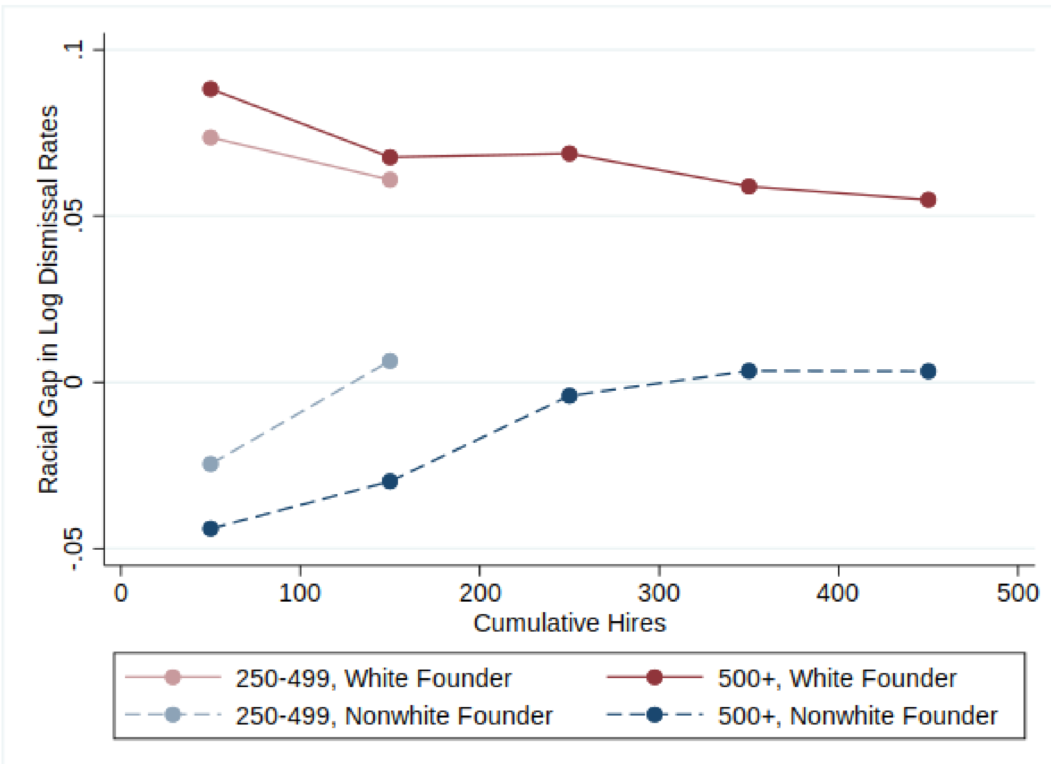
We emphasize a demand-side interpretation of the convergence patterns that we document: preferences over workers by race converge between firms with white and nonwhite founders. But there is also a supply-side explanation. Rather than firm preferences varying with founder race and cumulative hires, convergence 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.

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

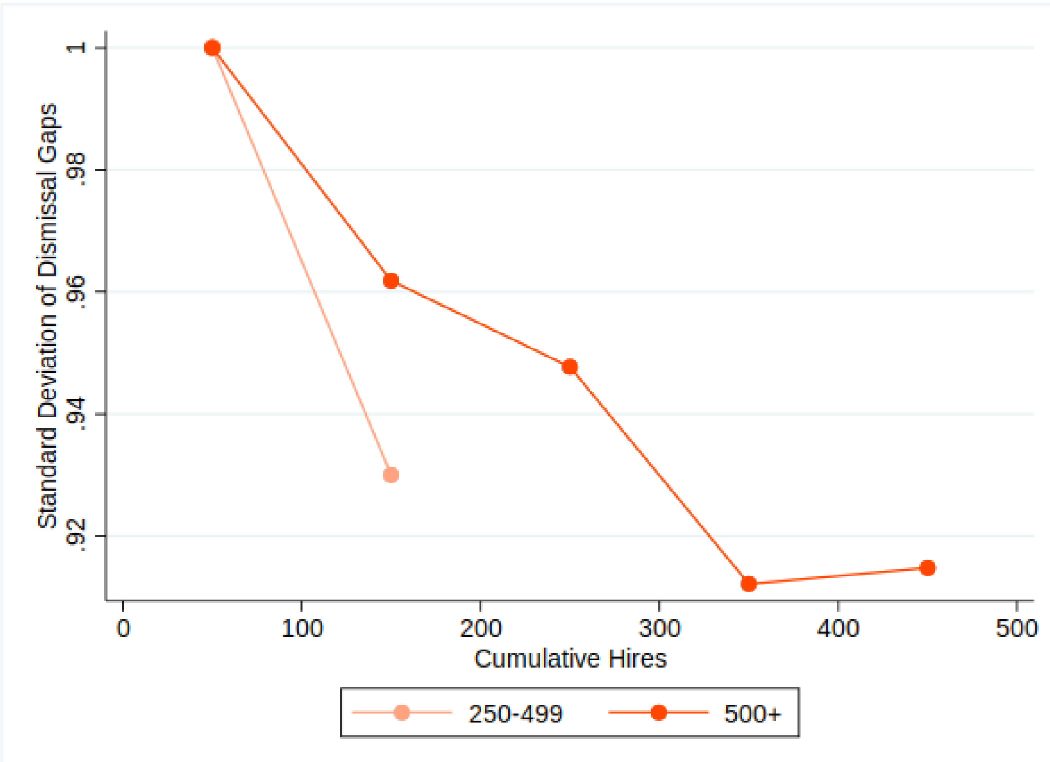
[19] The pattern is similar if we include all separations rather than limit to dismissals (see Appendix Figure C.2).

FIGURE 5
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 6
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 dismissal 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).

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

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

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.6 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 4.

5 Embedding co-racial hiring in a search model

The firm-level model studies hiring paths holding the pool of job seekers fixed. This section embeds the same co-racial hiring mechanism in a stationary search model and endogenizes that pool. The model asks how co-racial hiring affects group-specific unemployment rates once unemployed composition, vacancy filling rates, and firm composition are jointly determined. It also shows why average hiring profiles over the firm life cycle are informative about how co-racial hiring affects the direction of aggregate unemployment gaps. Appendix B.3 gives the formal environment, derivations, and proofs.

In the model, each (live) firm posts the same mass of vacancies per period. Hence, aggregate vacancy supply is fixed. We also abstract from wage determination: wages do not affect search, vacancy filling, or the allocation of vacancies across groups. The model therefore isolates the composition margin: co-racial hiring affects how a given flow of vacancies is directed across worker groups, not how many vacancies firms create or what wages they pay.[20]

Because each firm posts a fixed flow of vacancies, firms in the aggregate model are indexed by age rather than cumulative hires. The empirical life-cycle profiles, by contrast, are indexed by cumulative hires. Age and expected cumulative hires move together, but they are not identical: when targeted vacancies have different group-specific filling rates, expected hires per period depend on the composition of vacancies. Nonetheless, we use the model's age profiles to interpret the cumulative hire evidence.

[20]This separation matters because, if firms could choose how many vacancies to post, they could affect workforce composition not only by directing vacancies toward different groups, but also by choosing when to hire more or fewer workers. A firm could post more vacancies when the expected composition of hires moves it toward a desired incumbent mix, or fewer vacancies when hiring would dilute that mix. The firm's vacancy-posting policy would then depend on firm composition, so equilibrium would require tracking the full distribution of firm composition, size, and vacancy choices rather than simply the nonwhite share of targeted vacancies.

5.1 Environment and mechanism

There are two worker groups, white and nonwhite, indexed by $g \in \{W, N\}$. Labor force masses are fixed at L^W and L^N, and

$$\ell \equiv \frac{L^N}{L^W + L^N} \in (0, 1)$$

denotes the nonwhite labor force share. Workers search only while unemployed. Let U^g denote the unemployed mass and $u^g \equiv U^g/L^g$ the unemployment rate for group g. The nonwhite share among unemployed job seekers is

$$\gamma \equiv \frac{U^N}{U^N + U^W} = \frac{\ell u^N}{\ell u^N + (1-\ell)u^W}.$$

Firms are indexed by founder type $\tau \in \{W, N\}$ and age $a = 0, 1, 2, \ldots$. Entrants begin with no employees but inherit a founder-linked recruiting orientation,

$$\pi_0^W = 0, \qquad \pi_0^N = 1.$$

Existing matches separate at rate $\delta \in (0, 1)$, firms die at rate $\lambda \in (0, 1)$, and a fixed mass M of new firms enters each period. A share $\phi \in [0, 1]$ of entrants has nonwhite founders. Each (live) firm posts a fixed mass $\bar{v}$ of vacancies per period, so the stationary aggregate supply of vacancies is

$$\bar{V} = \frac{M}{\lambda}\bar{v}.$$

The share of a firm's vacancies targeted toward nonwhite job seekers is

$$p(\pi, \gamma) = (1-\rho)\gamma + \rho\pi, \qquad \rho \in [0, 1),$$

where $\rho = 0$ nests the case with no co-racial hiring force, and $\rho > 0$ makes recruiting depend on firm composition.

Vacancies are posted in two directed submarkets: one targeted to nonwhite job seekers and one targeted to white job seekers. Let V^g be vacancies targeted to group g, and let

$$\theta_g \equiv \frac{V^g}{U^g}$$

be group-specific tightness. The matching technology is

$$q(\theta) = m\theta^{-\alpha}, \qquad f(\theta) = \theta q(\theta) = m\theta^{1-\alpha}, \qquad \text{where } \alpha \in (0, 1),$$

where q is the vacancy-filling rate and f is the job-finding rate. We normalize tightness so that the balanced allocation with $P = \ell$ has $\theta_N = \theta_W = 1$. In that symmetric benchmark, $q(1) = f(1) = m$, so m is the common job-finding and vacancy-filling rate. The parameter α governs how sensitive the vacancy-filling rate is to how crowded a submarket is: when a group receives many targeted vacancies relative to its unemployment pool, those vacancies become harder to fill, and this adjustment is stronger when α is larger.

This structure creates two opposing margins. If nonwhite workers receive too small a share of targeted vacancies, then nonwhite tightness is low. For workers, this means lower job-finding rates and higher unemployment. For firms, however, low tightness makes nonwhite-targeted vacancies easier to fill. Hence under-targeting a group creates a market response that pushes firms back toward that group.

A firm with composition π obtains expected hires

$$h^N = \bar{v}p(\pi,\gamma)q(\theta_N), \qquad h^W = \bar{v}[1-p(\pi,\gamma)]q(\theta_W).$$

Hence, targeted recruiting and completed hiring need not coincide. A firm's completed hire share therefore depends both on where it aims its vacancies and on how easily those vacancies fill:

$$\widehat{p}(\pi,\gamma,\theta_N,\theta_W) = \frac{p(\pi,\gamma)q(\theta_N)}{p(\pi,\gamma)q(\theta_N) + [1-p(\pi,\gamma)]q(\theta_W)}.$$

5.2 Firm-side aggregation

Fix a candidate aggregate state $(\gamma,\theta_N,\theta_W)$ and write

$$r \equiv \frac{q(\theta_N)}{q(\theta_W)}$$

for the nonwhite-to-white vacancy-filling-rate ratio. As a firm ages, its workforce composition changes, and so does the share of vacancies it aims at nonwhite job seekers. Given (γ, r) and founder type τ, denote this sequence by $\{p_a^\tau\}_{a\geq 0}$. The stationary age-weighted targeted share for founder type τ is

$$P_\tau(\gamma,r) \equiv \lambda\sum_{a=0}^{\infty}(1-\lambda)^a p_a^\tau, \quad \text{where } \tau \in \{W,N\}.$$

Proposition 5.1 (Founder type targeting). *For each founder type τ, the map $(\gamma,r)\mapsto P_\tau(\gamma,r)$ is continuous and weakly increasing in both arguments. If $\rho>0$, then*

$$P_N(\gamma,r) > P_W(\gamma,r)$$

for every $\gamma\in(0,1)$ and $r>0$.

A more nonwhite unemployed pool raises nonwhite-targeted recruiting, and making nonwhite-targeted vacancies easier to fill has the same effect. Founder composition matters because nonwhite-founded firms direct a larger share of vacancies toward nonwhite job seekers than white-founded firms facing the same aggregate conditions.

5.3 Equilibrium characterization

Let

$$s^H \equiv \frac{H^N}{H^N + H^W}$$

denote the stationary nonwhite share of completed hires. The stationary unemployment rate for a group facing tightness θ is

$$u(\theta) = \frac{\bar{\delta}}{\bar{\delta} + f(\theta)} = \frac{\bar{\delta}}{\bar{\delta} + m\theta^{1-\alpha}}, \qquad \text{where } \bar{\delta} \equiv \frac{1 - (1-\lambda)(1-\delta)}{1-\lambda}.$$

The next result links the sign of the unemployment gap to the nonwhite share of completed hires.

Proposition 5.2 (Unemployment gaps, tightness, and aggregate hiring)**.** *In a stationary equilibrium,*

$$u^N > u^W \iff \theta_N < \theta_W \iff q(\theta_N) > q(\theta_W) \iff s^H < \ell.$$

Because the aggregate supply of vacancies is fixed, the natural fixed point object is the targeted nonwhite vacancy share

$$P \equiv \frac{V^N}{\bar{V}}.$$

A candidate value of P pins down group-specific tightness through vacancy-market clearing:

$$\theta_N(P)u(\theta_N(P)) = \frac{\bar{V}P}{L^N}, \qquad \theta_W(P)u(\theta_W(P)) = \frac{\bar{V}(1-P)}{L^W}.$$

The function $\theta \mapsto \theta u(\theta)$ is strictly increasing, so these equations uniquely determine $\theta_N(P)$ and $\theta_W(P)$. These tightness levels imply an unemployed pool share $\gamma(P)$ and a filling-rate ratio $r(P)$. Given $\gamma(P)$ and $r(P)$, firms of each founder type choose a nonwhite share of targeted vacancies. Weighting these choices by the founder composition gives the aggregate nonwhite share of targeted vacancies implied by firms' recruiting decisions:

$$\Psi(P;\phi) \equiv (1-\phi)P_W(\gamma(P), r(P)) + \phi P_N(\gamma(P), r(P)).$$

A stationary equilibrium is a fixed point of this map.

Proposition 5.3 (Fixed point characterization and uniqueness)**.** *For every founder share $\phi \in [0,1]$, stationary equilibria of the model correspond to fixed points of*

$$P = \Psi(P;\phi).$$

The map $\Psi(\cdot;\phi)$ is continuous and weakly decreasing on $(0,1)$, takes values in $(0,1)$, and the fixed point equation has a unique solution, denoted $P^*(\phi)$. Moreover, at the fixed point,

$$u^N > u^W \iff P^*(\phi) < \ell \iff s^H < \ell.$$

The proposition says that one aggregate object summarizes the equilibrium: the share of vacancies aimed at nonwhite job seekers. Given a candidate value of this share, the matching market determines group-specific tightness, unemployment composition, and filling rates. Firms then respond to those conditions through their recruiting rules. Equilibrium requires the vacancy share implied by firm behavior to coincide with the candidate share.

5.4 Founder composition and unemployment gaps

The natural benchmark is the allocation in which vacancies are targeted in proportion to the labor force. At $P = \ell$, the two groups face the same tightness, the unemployed pool has the same composition as the labor force, and filling rates are equal. If the founder mix also matches the labor force, $\phi = \ell$, there is no aggregate force pushing firms toward either group, and unemployment rates are equal.

To identify the founder share (ϕ) threshold, we evaluate the fixed point map at the allocation in which vacancies are targeted in proportion to the labor force, $P = \ell$. Then the two groups face the same tightness, the unemployed pool mirrors the labor force, and filling rates are equal across groups. Aggregating the firm-side dynamics across founder types gives

$$\Psi(\ell;\phi) = \ell + \rho D(\delta,\lambda,\rho)(\phi - \ell), \qquad D(\delta,\lambda,\rho) \in (0,1).$$

The attenuation factor D (defined explicitly in Appendix Lemma B.4) summarizes how entry, exit, and separation weaken the influence of the initial founder orientation as firms age. Hence, when $\phi < \ell$, firms collectively direct too few vacancies toward nonwhite job seekers relative to parity; when $\phi > \ell$, they direct too many. Since the fixed point is unique, this comparison determines the sign of the equilibrium unemployment gap.

The market response is stabilizing but self-limiting. When nonwhite workers are under-targeted, their unemployment share rises and nonwhite-targeted vacancies become easier to fill, both of which induce firms to target more nonwhite workers. But as this adjustment closes the unemployment gap, the scarcity signal weakens. At $P = \ell$, the signal disappears entirely: $\gamma = \ell$ and $r = 1$. If $\phi < \ell$, the founder-composition force remains at that point, so firms still choose a nonwhite vacancy share below ℓ. Therefore parity cannot be an equilibrium.

Proposition 5.4 (Founder share threshold and monotonicity)**.** *Fix $\rho \in (0,1)$. In the unique stationary equilibrium, $P^*(\phi)$ is strictly increasing in ϕ and*

$$\phi < \ell \Longrightarrow P^*(\phi) < \ell \Longrightarrow u^N > u^W$$

$$\phi = \ell \Longrightarrow P^*(\phi) = \ell \Longrightarrow u^N = u^W,$$

$$\phi > \ell \Longrightarrow P^*(\phi) > \ell \Longrightarrow u^N < u^W.$$

Moreover, $u^N - u^W$ is strictly decreasing in ϕ, and the primitive threshold for unemployment parity *is exactly $\phi = \ell$.*

The result gives a simple threshold rule. Co-racial hiring disadvantages nonwhite workers in unemployment exactly when nonwhite founders are underrepresented relative to the nonwhite labor-force share. When nonwhite founders are overrepresented, the same mechanism works in the opposite direction.

5.5 Life-cycle profiles and aggregate gaps

The threshold result also gives an interpretation to hiring profiles over firm age. In the data, one key object is whether the average firm's nonwhite share of hires rises or falls as the firm accumulates hires. In the model, founder type affects early recruiting, while firms converge to a common steady-state nonwhite share of hires.[21] In the equilibrium associated with founder share ϕ, let $\pi^*(\phi)$ denote this common target and let $s_0^H(\phi)$ denote the average initial nonwhite share of hires.

Proposition 5.5 (Life-cycle profiles and aggregate gaps)**.** *Fix $\rho \in (0, 1)$ and suppose the equilibrium satisfies the contraction condition in (2).[22] Then*

$$\operatorname{sign}\bigl(\pi^*(\phi) - s_0^H(\phi)\bigr) = \operatorname{sign}(u^N - u^W) = \operatorname{sign}(\ell - \phi).$$

Hence, the average initial nonwhite hire share is below the common target if and only if $\phi < \ell$, equal to the common target if and only if $\phi = \ell$, and above the common target if and only if $\phi > \ell$.

This proposition connects the aggregate model to the life-cycle evidence. We interpret the plateau value in the empirical life-cycle profile as the empirical analog of the steady-state nonwhite hire share, $\pi^*(\phi)$. Since entrant firms begin below this plateau on average, the empirical counterpart of $\pi^*(\phi) - s_0^H(\phi)$ is positive. Under the maintained co-racial hiring mechanism and the equilibrium contraction condition, Proposition 5.5 therefore implies $\phi < \ell$. Proposition 5.4 then implies that co-racial hiring raises nonwhite unemployment relative to white unemployment.

Conversely, if the average entrant began above the steady-state nonwhite hire share, the model would imply $\phi > \ell$ and the opposite unemployment ranking. In the model, a flat average profile corresponds either to the symmetric benchmark in which the founder mix matches the labor force composition, $\phi = \ell$, or to the case $\rho = 0$, in which incumbent composition does not affect future recruiting.

Appendix B.4 studies a broader affine class,

$$p(\pi, \gamma) = a(\gamma) + b(\gamma)\pi,$$

[21] See Appendix Lemma B.5.

[22] In the aggregate equilibrium associated with founder share ϕ, this condition is

$$\rho \max\left\{\frac{q(\theta_N(P^*(\phi)))}{q(\theta_W(P^*(\phi)))}, \frac{q(\theta_W(P^*(\phi)))}{q(\theta_N(P^*(\phi)))}\right\} < 1.$$

This is the same contraction condition as (2), evaluated at the equilibrium filling rates.

which can naturally represent screening discrimination, taste-based discrimination, and production complementarities. The benchmark linear rule is a special case. In the more general affine setting, many firm-level implications continue to hold; in particular, when nonwhites make up at most half of the labor force, a rising average life-cycle profile still implies $u^N > u^W$.

5.6 Comparative statics in other key parameters are ambiguous

The model gives a sharp prediction about the direction of the unemployment gap, but not about how its magnitude changes with every parameter. Parameters such as the matching elasticity α, the strength of co-racial hiring ρ, and the firm exit rate λ each affect equilibrium through multiple channels. For example, a stronger market-clearing response can make vacancies aimed at an under-targeted group easier to fill, which tends to narrow unemployment gaps, but it also changes the composition of the unemployed pool to which firms' recruiting is partly anchored. Similarly, stronger co-racial hiring and higher firm turnover can amplify founder-driven differences in recruiting, while also inducing offsetting changes in unemployment composition and vacancy-filling rates. We therefore emphasize the sign result in Proposition 5.4, which is robust to these equilibrium adjustments, rather than general comparative statics for the size of the gap.

6 Discussion

This paper studies the firm-level and aggregate consequences of co-racial hiring. We develop a model in which recruiting is shaped by both the racial composition of a firm's incumbent workforce and the composition of external job seekers. Consistent with the model, we document a striking empirical pattern: as a firm's cumulative hires increase, so does the probability that its next hire is nonwhite. This pattern is not explained by changes in observable job characteristics. We also find that firms with white and nonwhite founders are more likely to hire workers of the founder's own race, and that these differences shrink as cumulative hires increase. Dismissals follow a similar pattern: firms are less likely to dismiss recent hires of the same race as the founder, and racial differences in dismissal rates also decline with cumulative hires.

We then embed co-racial hiring in a search model with group-specific matching. In the model, the firm-level effects of co-racial hiring may be partly offset by general equilibrium forces: when one group is relatively under-targeted, its share of unemployed job seekers rises and vacancies directed toward that group become easier to fill. Through the lens of the model, our stylized fact that later hires are more likely to be nonwhite than earlier hires implies that co-racial hiring increases nonwhite unemployment relative to white unemployment.

While our main analysis focuses on Brazil, publicly available U.S. data provide suggestive evidence of a similar rising life-cycle profile for Black hires. This comparison is relevant because racial differences in employment and entrepreneurship are also large in the United States. In Appendix D, we use Quarterly Workforce Indicators (QWI), Census tabulations derived from Longitudinal Employer-Household Dynamics (LEHD) data, which report new hires by worker race, quarter, firm age, geography, and industry. We use these tabulations to construct aggregate firm-entry cohorts: within each geography-by-industry cell, we group firm age bins by implied entry year and ask how the Black share of new hires changes as those cohorts age. We implement the exercise at both the state and county levels, limiting the county analysis to the 100 most populous counties in 2000. In both analyses, the Black share of new hires rises as cohorts move from ages 0–1 to ages 2–3 and 4–5. The public QWI tabulations do not allow us to identify individual firms or measure cumulative hires within firms, so the evidence is necessarily indirect. Still, the direction of the profile is consistent with the central life-cycle pattern we document in the Brazilian formal sector.[23]

Our findings suggest several ways policy can reduce or contribute to racial inequality in labor demand. First, they provide a rationale for affirmative action policies. Over the firm life cycle, founder-linked differences in the racial composition of hires fade as firms accumulate hires. But this adjustment is slow, and few firms reach the scale at which founder race no longer predicts the racial composition of hires. Temporary affirmative action policies could accelerate this adjustment by encouraging firms to hire workers from groups that are underrepresented within the firm relative to the relevant external labor market. Such policies may have short-run costs—for example, higher dismissal rates—but our results suggest that they may also produce persistent reductions in racial inequality in labor demand (Miller, 2017). Although affirmative action policies often exclude small firms, our findings suggest that they may be especially effective at small or young firms.

Second, our results point to a direct link between racial differences in entrepreneurship and labor demand. This connection helps motivate policies that encourage entrepreneurship among underrepresented groups.

Third, our findings suggest that market frictions affecting the size distribution of firms may also shape racial inequality in the labor market (Restuccia and Rogerson, 2017). For example, if small but productive firms are unable to grow to efficient scale because of resource misallocation, they are also less likely to reach the point at which their workforces reflect the broader labor market. By this logic, the aggregate costs of misallocation may fall disproportionately on groups that are underrepresented among entrepreneurs.

[23]The finding also complements earlier U.S. evidence that the Black share of employment rises with employer size overall, except among Black-run businesses, where that pattern is reversed (Holzer, 1998; Sorkin, 2018).

References

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.

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.

Black, Dan A., "Discrimination in an Equilibrium Search Model," *Journal of Labor Economics*, 1995, *13* (2), 309–334.

Bossler, Mario, Alexander Mosthaf, and Thorsten Schank, "Are Female Managers More Likely to Hire More Female Managers? Evidence from Germany," *ILR Review*, may 2020, *73* (3), 676–704.

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.

Bowlus, Audra J. and Zvi Eckstein, "Discrimination and Skill Differences in an Equilibrium Search Model," *International Economic Review*, November 2002, *43* (4), 1309–1345.

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.

Burks, Stephen V., Bo Cowgill, Mitchell Hoffman, and Michael Housman, "The Value of Hiring through Employee Referrals," *Quarterly Journal of Economics*, May 2015, *130* (2), 805–839.

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.

Cullen, Zoë and Ricardo Perez-Truglia, "The Old Boys' Club: Schmoozing and the Gender Gap," *American Economic Review*, july 2023, *113* (7), 1703–1740.

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.

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

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, François, 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*, 2021, *111* (10), 3418–3457.

Giuliano, Laura, David I. Levine, and Jonathan Leonard, "Manager Race and the Race of New Hires," *Journal of Labor Economics*, October 2009, *27* (4), 589–631.

_ , _ , and _ , "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.

Glover, Dylan, Amanda Pallais, and William Pariente, "Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores," *The Quarterly Journal of Economics*, August 2017, *132* (3), 1219–1260.

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.

Hensvik, Lena E., "Manager Impartiality: Worker-Firm Matching and the Gender Wage Gap," *ILR Review*, 2014, *67* (2), 395–421.

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

Hjort, Jonas, "Ethnic Divisions and Production in Firms," *Quarterly Journal of Economics*, November 2014, *129* (4), 1899–1946.

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

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.

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

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.

Montgomery, James D, "Social Networks and Labor Market Outcomes: Toward an Economic Analysis," *American Economic Review*, 1991, *81*, 1408.

Pallais, Amanda and Emily Glassberg Sands, "Why the Referential Treatment? Evidence from Field Experiments on Referrals," *Journal of Political Economy*, December 2016, *124* (6), 1793–1828.

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.

Ritter, Joseph A. and Lowell J. Taylor, "Racial Disparity in Unemployment," *Review of*

Economics and Statistics, February 2011, *93* (1), 30–42.

Rosén, Åsa, "Search, Bargaining, and Employer Discrimination," *Journal of Labor Economics*, October 2003, *21* (4), 807–830.

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.

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

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.

Stoll, Michael A., Steven Raphael, and Harry J. Holzer, "Black Job Applicants and the Hiring Officer's Race," *ILR Review*, 2004, *57* (2), 267–287.

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

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

Online Appendix:

Co-Racial Hiring:
Firm Dynamics and Aggregate Consequences

Conrad Miller

Ian Schmutte

May 2026

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 three 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 `Job` 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 the 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.

B Theory appendix

We collect the formal proofs for the firm-level and macro model sections.

B.1 Firm-level dynamics with fixed filling rates

Fix $\gamma \in [0,1]$ and $q_N, q_W > 0$. Let

$$F_{q_N,q_W}(p) \equiv \frac{p\,q_N}{p\,q_N + (1-p)\,q_W}, \qquad p(\pi,\gamma) = (1-\rho)\gamma + \rho\pi, \qquad \widehat{p}(\pi,\gamma;q_N,q_W) = F_{q_N,q_W}(p(\pi,\gamma)).$$

Then

$$\Pr(Y_{k+1} = 1 \mid \pi_k, \gamma, q_N, q_W) = \widehat{p}(\pi_k, \gamma; q_N, q_W)$$

and

$$\mathbb{E}[\pi_{k+1} \mid \pi_k, \gamma, q_N, q_W] = \pi_k + \frac{\widehat{p}(\pi_k, \gamma; q_N, q_W) - \pi_k}{k+1}. \tag{B.1}$$

Completed hiring with unequal filling rates. For every $\pi, \gamma \in [0,1]$ and $q_N, q_W > 0$,

$$\widehat{p}(\pi,\gamma;q_N,q_W) - p(\pi,\gamma) = \frac{p(\pi,\gamma)[1-p(\pi,\gamma)](q_N - q_W)}{p(\pi,\gamma)\,q_N + [1-p(\pi,\gamma)]\,q_W}. \tag{B.2}$$

Hence completed hires tilt toward the group with the higher filling rate, except at corner recruiting shares. To see this, let $p = p(\pi,\gamma)$ and

$$D = pq_N + (1-p)q_W.$$

Then

$$\widehat{p}(\pi,\gamma;q_N,q_W) - p = \frac{pq_N}{D} - p = \frac{pq_N - p[pq_N + (1-p)q_W]}{D}.$$

The numerator is

$$pq_N - p^2 q_N - p(1-p)q_W = p(1-p)(q_N - q_W),$$

which gives (B.2). Since the denominator is positive, the sign follows from $q_N - q_W$ whenever $p(\pi,\gamma) \in (0,1)$.

Proposition B.1 (Common target under fixed filling rates)**.** *Suppose the filling-rate-adjusted contraction bound satisfies*

$$\kappa \equiv \rho\,\frac{\max\{q_N, q_W\}}{\min\{q_N, q_W\}} < 1.$$

Then $\widehat{p}(\cdot,\gamma;q_N,q_W)$ is a contraction on $[0,1]$, so it has a unique fixed point

$$\pi^*(\gamma,q_N,q_W) = \widehat{p}(\pi^*(\gamma,q_N,q_W),\gamma;q_N,q_W).$$

Moreover,

$$\mathbb{E}[\pi_{k+1} - \pi_k \mid \pi_k, \gamma, q_N, q_W] = \frac{\widehat{p}(\pi_k,\gamma;q_N,q_W) - \pi_k}{k+1} \gtrless 0 \qquad \Longleftrightarrow \qquad \pi_k \lessgtr \pi^*(\gamma,q_N,q_W). \tag{B.3}$$

Founder type changes the path through the initial condition, not the target.

Proof. Because $p(\pi, \gamma) = (1-\rho)\gamma + \rho\pi$,

$$|p(\pi, \gamma) - p(\tilde{\pi}, \gamma)| = \rho|\pi - \tilde{\pi}|.$$

Also,

$$F'_{q_N, q_W}(p) = \frac{q_N q_W}{[pq_N + (1-p)q_W]^2} \leq \frac{\max\{q_N, q_W\}}{\min\{q_N, q_W\}} \qquad \text{for all } p \in [0,1].$$

Hence

$$|\widehat{p}(\pi, \gamma; q_N, q_W) - \widehat{p}(\tilde{\pi}, \gamma; q_N, q_W)| \leq \kappa|\pi - \tilde{\pi}|.$$

The contraction mapping theorem gives the unique fixed point.

Let $h(\pi) \equiv \widehat{p}(\pi, \gamma; q_N, q_W) - \pi$. For $\pi > \tilde{\pi}$,

$$h(\pi) - h(\tilde{\pi}) = \widehat{p}(\pi, \gamma; q_N, q_W) - \widehat{p}(\tilde{\pi}, \gamma; q_N, q_W) - (\pi - \tilde{\pi}) < 0,$$

so h is strictly decreasing. Since $h(\pi^*) = 0$, the sign of $h(\pi_k)$ is the sign of $\pi^* - \pi_k$. Combining this with (B.1) gives (B.3). ☐

Proposition B.2 (Firm-level implications)**.** *Under the hypotheses of Proposition B.1, suppose also that $\rho > 0$ and $\pi_0^N > \pi_0^W$. Let $\{\pi_k^N\}_{k\geq 0}$ and $\{\pi_k^W\}_{k\geq 0}$ denote the random sequences of firm nonwhite employment shares for two firms facing the same (γ, q_N, q_W), started from π_0^N and π_0^W, respectively. Then:*

(i) the nonwhite-founder firm has a higher first-hire probability;

(ii) the expected composition gap shrinks with cumulative hires and converges to zero; the expected hire-probability gap also converges to zero and is bounded above by a decreasing sequence;

(iii) dispersion in conditional mean composition and conditional mean hire probabilities across initial orientations is bounded by terms that converge to zero;

(iv) conditional on current composition, founder type and firm age have no additional effect on the next-hire probability.

Proof. The map $\pi \mapsto \widehat{p}(\pi, \gamma; q_N, q_W)$ is strictly increasing because $\rho > 0$ and F_{q_N, q_W} is strictly increasing. Hence the first-hire comparison follows from $\pi_0^N > \pi_0^W$.

Define

$$A_0(\kappa) \equiv 1, \qquad A_k(\kappa) \equiv \prod_{m=0}^{k-1}\left(1 - \frac{1-\kappa}{m+1}\right) \quad \text{for } k \geq 1.$$

Since $\kappa < 1$, $A_k(\kappa) \to 0$. Couple the two firms using a common sequence $\{U_{k+1}\}_{k\geq 0}$ of i.i.d. Uniform$[0,1]$ draws and set

$$Y_{k+1}^\tau = \mathbf{1}\{U_{k+1} \leq \widehat{p}(\pi_k^\tau, \gamma; q_N, q_W)\}, \qquad \tau \in \{N, W\}.$$

Monotonicity gives $\pi_k^N \geq \pi_k^W$ almost surely for every k. Let $\Delta_k \equiv \pi_k^N - \pi_k^W$. Using the κ-Lipschitz bound,

$$\mathbb{E}[\Delta_{k+1} \mid \pi_k^N, \pi_k^W] \leq \left(1 - \frac{1-\kappa}{k+1}\right) \Delta_k.$$

Iterating gives

$$0 \leq \mathbb{E}[\pi_k^N] - \mathbb{E}[\pi_k^W] \leq A_k(\kappa)(\pi_0^N - \pi_0^W),$$

and therefore the expected composition gap converges to zero. Applying the same Lipschitz bound once more gives

$$0 \leq \mathbb{E}[\widehat{p}(\pi_k^N, \gamma; q_N, q_W)] - \mathbb{E}[\widehat{p}(\pi_k^W, \gamma; q_N, q_W)] \leq \kappa A_k(\kappa)(\pi_0^N - \pi_0^W),$$

so the expected hire-probability gap converges to zero and is bounded above by the decreasing sequence $\kappa A_k(\kappa)(\pi_0^N - \pi_0^W)$.

For dispersion, let Π_0 be a random initial orientation and define

$$\mu_k(\pi_0) \equiv \mathbb{E}[\pi_k \mid \pi_0, \gamma, q_N, q_W], \qquad \eta_k(\pi_0) \equiv \mathbb{E}[\widehat{p}(\pi_k, \gamma; q_N, q_W) \mid \pi_0, \gamma, q_N, q_W].$$

The same coupling argument gives

$$|\mu_k(\pi_0) - \mu_k(\tilde{\pi}_0)| \leq A_k(\kappa)|\pi_0 - \tilde{\pi}_0|, \qquad |\eta_k(\pi_0) - \eta_k(\tilde{\pi}_0)| \leq \kappa A_k(\kappa)|\pi_0 - \tilde{\pi}_0|.$$

Using $\mathrm{Var}(X) = \frac{1}{2}\mathbb{E}[(X - X')^2]$ with an independent copy gives

$$\mathrm{Var}(\mu_k(\Pi_0)) \leq A_k(\kappa)^2 \, \mathrm{Var}(\Pi_0), \qquad \mathrm{Var}(\eta_k(\Pi_0)) \leq \kappa^2 A_k(\kappa)^2 \, \mathrm{Var}(\Pi_0).$$

Finally,

$$\Pr(Y_{k+1} = 1 \mid \pi_k, \gamma, q_N, q_W, \tau, k) = \widehat{p}(\pi_k, \gamma; q_N, q_W),$$

so founder type and age matter for the next hire only through current composition. □

Equal filling rates. If $q_N = q_W = q > 0$, then

$$\widehat{p}(\pi, \gamma; q, q) = p(\pi, \gamma) = (1-\rho)\gamma + \rho\pi$$

and the common target is γ. The bounds above are exact:

$$\mathbb{E}[\pi_k \mid \tau, \gamma] - \gamma = A_k(\rho)(\pi_0^\tau - \gamma),$$

and

$$\mathbb{E}[p(\pi_k, \gamma) \mid N, \gamma] - \mathbb{E}[p(\pi_k, \gamma) \mid W, \gamma] = \rho A_k(\rho)(\pi_0^N - \pi_0^W),$$

where

$$A_0(\rho) = 1, \qquad A_k(\rho) = \prod_{m=0}^{k-1} \left(1 - \frac{1-\rho}{m+1}\right) \quad \text{for } k \geq 1.$$

B.2 Microfoundations and the affine class

The benchmark rule in the text,

$$p(\pi, \gamma) = (1-\rho)\gamma + \rho\pi,$$

is the simplest member of a broader affine class. For a given vacancy, suppose the relevant incumbent or decision-maker is nonwhite with probability π. Let $p_N(\gamma)$ be the nonwhite share of effective recruiting when that focal person is nonwhite and $p_W(\gamma)$ the analogous share when that focal person is white. Then

$$p(\pi, \gamma) = \pi p_N(\gamma) + (1-\pi) p_W(\gamma) = a(\gamma) + b(\gamma)\pi,$$

where $a(\gamma) = p_W(\gamma)$ and $b(\gamma) = p_N(\gamma) - p_W(\gamma)$. Here, "recruiting" refers to the object affected by the mechanism: targeted search, screening, or productive matches. Throughout this appendix we assume $0 \leq p_W(\gamma) \leq p_N(\gamma) \leq 1$ and that both functions are weakly increasing in γ.

The linear benchmark is the special case

$$p_N(\gamma) = (1-\rho)\gamma + \rho, \qquad p_W(\gamma) = (1-\rho)\gamma.$$

Referral hiring. If a fraction ω of recruiting comes through referrals and a fraction $1-\omega$ comes through the external market, and if referral candidates have nonwhite share

$$(1-\eta)\gamma + \eta\pi, \qquad \eta \in [0,1],$$

then

$$p(\pi, \gamma) = \omega[(1-\eta)\gamma + \eta\pi] + (1-\omega)\gamma = (1-\rho)\gamma + \rho\pi, \qquad \rho \equiv \omega\eta.$$

Segregated referral networks therefore deliver the linear rule.

Screening discrimination. Suppose the applicant pool has nonwhite share γ. A focal screener evaluates same-group candidates with yield ξ_H and other-group candidates with yield ξ_L, where $\xi_H > \xi_L > 0$. If the hire is drawn in proportion to the arrival of acceptable candidates, then

$$p_N(\gamma) = \frac{\xi_H \gamma}{\xi_H \gamma + \xi_L(1-\gamma)}, \qquad p_W(\gamma) = \frac{\xi_L \gamma}{\xi_L \gamma + \xi_H(1-\gamma)}.$$

Screening discrimination therefore yields an affine rule, though generally not the linear benchmark.

Taste-based discrimination. For each vacancy, suppose a focal decision-maker is drawn from the firm's incumbent workforce. The focal decision-maker is nonwhite with probability π and white with probability $1-\pi$. The candidate pool has nonwhite share γ and white share $1-\gamma$.

Let the decision-maker's group be $h \in \{N, W\}$ and candidate i's group be $g_i \in \{N, W\}$. The decision-maker assigns candidate i perceived value

$$U_i(h) = x_i + \beta \mathbf{1}\{g_i = h\} + \varepsilon_i, \qquad \beta > 0,$$

where x_i is candidate quality, β is a same-group taste premium, and ε_i is an idiosyncratic evaluation shock. Assume that the distribution of x_i is the same across groups, so that absent the taste premium the selected candidate is nonwhite with probability γ. Suppose the idiosyncratic shocks are independent type-I extreme value, and the decision-maker chooses the candidate with the highest perceived value.

Then, conditional on a nonwhite decision-maker, each nonwhite candidate receives the taste premium, so the probability that the selected candidate is nonwhite is

$$p_N(\gamma) = \frac{e^{\beta}\gamma}{e^{\beta}\gamma + (1-\gamma)}.$$

Conditional on a white decision-maker, each white candidate receives the taste premium, so

$$p_W(\gamma) = \frac{\gamma}{\gamma + e^{\beta}(1-\gamma)}.$$

Averaging over the race of the focal decision-maker gives

$$p(\pi, \gamma) = \pi p_N(\gamma) + (1-\pi) p_W(\gamma) = p_W(\gamma) + [p_N(\gamma) - p_W(\gamma)]\pi.$$

Thus taste-based discrimination generates an affine rule in incumbent composition.

The extreme-value assumption gives the logit formulas in closed form; the affine dependence on π follows more generally from averaging over the race of the focal decision-maker.

Production complementarities. A similar structure applies when the same-group mechanism operates through productivity rather than preferences. Suppose the new hire will work mainly with a focal incumbent, such as a supervisor, trainer, or close teammate. The focal incumbent is nonwhite with probability π and white with probability $1-\pi$. The candidate pool has nonwhite share γ and white share $1-\gamma$.

Let the focal incumbent's group be $h \in \{N, W\}$ and candidate i's group be $g_i \in \{N, W\}$. The firm assigns candidate i surplus

$$S_i(h) = x_i + \delta \mathbf{1}\{g_i = h\} + \varepsilon_i, \qquad \delta > 0,$$

where x_i is candidate quality, δ is a same-group productivity premium, and ε_i is an idiosyncratic match component. The premium δ represents more effective team production when the new hire and the focal incumbent belong to the same group. Assume that the distribution of x_i is the same across groups, so that absent the productivity premium the selected candidate is nonwhite with probability γ. Suppose the idiosyncratic match terms are independent type-I extreme value shocks, and the firm chooses the candidate with the highest surplus.

Then, conditional on a nonwhite focal incumbent, each nonwhite candidate receives the productivity premium, so the probability that the selected candidate is nonwhite is

$$p_N(\gamma) = \frac{e^{\delta}\gamma}{e^{\delta}\gamma + (1-\gamma)}.$$

Conditional on a white focal incumbent, each white candidate receives the productivity premium, so

$$p_W(\gamma) = \frac{\gamma}{\gamma + e^{\delta}(1-\gamma)}.$$

Averaging over the focal incumbent gives

$$p(\pi,\gamma) = \pi p_N(\gamma) + (1-\pi)p_W(\gamma) = p_W(\gamma) + [p_N(\gamma) - p_W(\gamma)]\pi.$$

Thus focal-incumbent production complementarities generate an affine rule in incumbent composition.

The three same-group mechanisms above therefore fit the same odds-premium class

$$p_N(\gamma) = \frac{\kappa\gamma}{1-\gamma+\kappa\gamma}, \qquad p_W(\gamma) = \frac{\gamma}{\kappa(1-\gamma)+\gamma}, \qquad \kappa > 1,$$

where κ is the screening yield ratio, e^{β}, or e^{δ}. When filling rates are equal, the common firm-level target under the affine rule is

$$\bar{\pi}(\gamma) \equiv \frac{a(\gamma)}{1-b(\gamma)} = \frac{p_W(\gamma)}{1-p_N(\gamma)+p_W(\gamma)}.$$

In the linear benchmark, $\bar{\pi}(\gamma) = \gamma$. Under the odds-premium class,

$$\bar{\pi}(\gamma) = \frac{\gamma[1+(\kappa-1)\gamma]}{\kappa - 2(\kappa-1)\gamma(1-\gamma)},$$

so $\bar{\pi}(\gamma) \leq \gamma$ whenever $\gamma \leq 1/2$. This is the case used in the affine macro discussion below.

B.3 Macro results

We next give the formal details for the macro analysis in section 5.

B.3.1 Formal environment and equilibrium

In each directed submarket, job-finding and vacancy-filling rates are given by:

$$q(\theta) = m\theta^{-\alpha}, \qquad f(\theta) = m\theta^{1-\alpha}, \qquad \alpha \in (0,1), \quad m > 0.$$

We choose units of tightness so that the symmetric benchmark has $\theta = 1$. Under this convention, $q(1) = f(1) = m$, so m is the common benchmark matching rate. The parameter α controls how quickly vacancies become harder to fill as targeted vacancies rise relative to the group's unemployment pool.

The nonwhite share of the unemployed pool is

$$\gamma \equiv \frac{U^N}{U^N + U^W} = \frac{\ell u^N}{\ell u^N + (1-\ell)u^W}.$$

Entrants inherit founder-specific initial orientation,

$$n_0^W = n_0^N = 0, \qquad \pi_0^W = 0, \qquad \pi_0^N = 1,$$

and entrant masses are

$$M_N = \phi M, \qquad M_W = (1-\phi)M.$$

Each live firm posts a fixed mass $\bar{v} > 0$ of vacancies, so total posted vacancies are

$$\bar{V} = \frac{M}{\lambda}\bar{v}.$$

The recruiting rule is

$$p(\pi, \gamma) = (1-\rho)\gamma + \rho\pi, \qquad \rho \in (0, 1),$$

and group-specific tightness is

$$\theta_g \equiv \frac{V^g}{U^g}, \qquad g \in \{W, N\}.$$

Given $\Gamma = (\gamma, \theta_N, \theta_W)$, the age profile for founder type τ is defined recursively by

$$p_a^\tau = (1-\rho)\gamma + \rho\pi_a^\tau, \qquad h_a^{\tau,N} = \bar{v}\, p_a^\tau q(\theta_N), \qquad h_a^{\tau,W} = \bar{v}\,[1 - p_a^\tau]q(\theta_W), \tag{B.4}$$

$$n_{a+1}^\tau = (1-\delta)n_a^\tau + h_a^{\tau,N} + h_a^{\tau,W}, \tag{B.5}$$

$$\pi_{a+1}^\tau = \begin{cases} \dfrac{(1-\delta)n_a^\tau \pi_a^\tau + h_a^{\tau,N}}{n_{a+1}^\tau}, & n_{a+1}^\tau > 0, \\ \pi_a^\tau, & n_{a+1}^\tau = 0. \end{cases} \tag{B.6}$$

The stationary mass of age-a firms with founder type τ is

$$\mu_a^\tau = (1-\lambda)^a M_\tau.$$

Aggregate employment, unemployment, targeted vacancies, and hires are

$$E^N = \sum_{\tau \in \{W,N\}} \sum_{a=0}^{\infty} \mu_a^\tau n_a^\tau \pi_a^\tau, \qquad E^W = \sum_{\tau \in \{W,N\}} \sum_{a=0}^{\infty} \mu_a^\tau n_a^\tau (1 - \pi_a^\tau),$$

$$U^N = L^N - E^N, \qquad U^W = L^W - E^W,$$

$$V^N = \bar{v} \sum_{\tau \in \{W,N\}} \sum_{a=0}^{\infty} \mu_a^\tau p_a^\tau, \qquad V^W = \bar{v} \sum_{\tau \in \{W,N\}} \sum_{a=0}^{\infty} \mu_a^\tau [1 - p_a^\tau],$$

and

$$H^N = \sum_{\tau \in \{W,N\}} \sum_{a=0}^{\infty} \mu_a^\tau h_a^{\tau,N}, \qquad H^W = \sum_{\tau \in \{W,N\}} \sum_{a=0}^{\infty} \mu_a^\tau h_a^{\tau,W}.$$

A stationary equilibrium consists of aggregate conditions and cohort paths satisfying the cohort recursion, the stationary age distribution, aggregate accounting for employment and unemployment, and matching-market clearing in both directed submarkets:

$$H^N = U^N f(\theta_N) = V^N q(\theta_N), \qquad H^W = U^W f(\theta_W) = V^W q(\theta_W).$$

B.3.2 Preliminary results

Define

$$\bar{\delta} \equiv \frac{1 - (1-\lambda)(1-\delta)}{1-\lambda}$$

and

$$u(\theta) \equiv \frac{\bar{\delta}}{\bar{\delta} + m\theta^{1-\alpha}}.$$

The symmetric benchmark has

$$\theta_N(\ell) = \theta_W(\ell) = 1.$$

Since $q(1) = f(1) = m$, this benchmark pins down

$$m = \bar{\delta}\left(\frac{L^W + L^N}{\bar{V}} - 1\right),$$

which is positive when $\bar{V} < L^W + L^N$.

Lemma B.3 (Monotone vacancy supply inversion)**.** *Let*

$$Z(\theta) \equiv \theta u(\theta) = \frac{\bar{\delta}\theta}{\bar{\delta} + m\theta^{1-\alpha}}.$$

Then Z is strictly increasing on $[0, \infty)$. Therefore, for each $P \in [0, 1]$, the equations

$$Z(\theta_N(P)) = \frac{\bar{V}P}{L^N}, \qquad Z(\theta_W(P)) = \frac{\bar{V}(1-P)}{L^W}$$

admit unique solutions. Moreover,

$$P < \ell \iff \theta_N(P) < \theta_W(P),$$

with equality at $P = \ell$ and the reverse inequality when $P > \ell$.

Proof. Direct differentiation gives

$$Z'(\theta) = \frac{\bar{\delta}[\bar{\delta} + \alpha m\theta^{1-\alpha}]}{[\bar{\delta} + m\theta^{1-\alpha}]^2} > 0.$$

Since $Z(0) = 0$ and $Z(\theta) \to \infty$ as $\theta \to \infty$, the inverse exists on $[0, \infty)$. Finally,

$$P < \ell \iff \frac{P}{L^N} < \frac{1-P}{L^W} \iff \frac{\bar{V}P}{L^N} < \frac{\bar{V}(1-P)}{L^W},$$

and strict monotonicity of Z gives the tightness comparison. The equality and $P > \ell$ cases are analogous. $\square$

Proof of Proposition 5.1. Fix $\gamma \in (0,1)$ and $r > 0$. Scale workforce components by the white filling rate and write

$$X_a^\tau = (x_a^{\tau,N}, x_a^{\tau,W}), \qquad x_a^{\tau,N} \equiv \frac{n_a^\tau \pi_a^\tau}{\bar{v}q_W}, \qquad x_a^{\tau,W} \equiv \frac{n_a^\tau(1-\pi_a^\tau)}{\bar{v}q_W}.$$

At age zero,

$$p_0^W = (1-\rho)\gamma, \qquad p_0^N = (1-\rho)\gamma + \rho,$$

so $p_0^W < p_0^N$ and

$$X_1^\tau = (rp_0^\tau, 1 - p_0^\tau).$$

For $a \geq 1$, define

$$\pi(X) \equiv \frac{x^N}{x^N + x^W}$$

and

$$T_{\gamma,r}(X) \equiv \Big((1-\delta)x^N + r[(1-\rho)\gamma + \rho\pi(X)],\ (1-\delta)x^W + 1 - [(1-\rho)\gamma + \rho\pi(X)]\Big).$$

Then $X_{a+1}^\tau = T_{\gamma,r}(X_a^\tau)$. If both components of X are nonnegative and at least one component is positive, then $T_{\gamma,r}(X)$ has strictly positive components.

Order states by

$$X \preceq Y \quad \iff \quad x^N \leq y^N \quad \text{and} \quad x^W \geq y^W,$$

and write $X \prec Y$ when one inequality is strict. This order means that Y has weakly more nonwhite employment and weakly less white employment. If $X \prec Y$, then $\pi(X) < \pi(Y)$ on strictly positive states. The transition is strictly order preserving:

$$T_{\gamma,r}^N(Y) - T_{\gamma,r}^N(X) = (1-\delta)(y^N - x^N) + r\rho[\pi(Y) - \pi(X)] > 0,$$

$$T_{\gamma,r}^W(X) - T_{\gamma,r}^W(Y) = (1-\delta)(x^W - y^W) + \rho[\pi(Y) - \pi(X)] > 0.$$

Since $X_1^W \prec X_1^N$, induction gives $X_a^W \prec X_a^N$ and hence $p_a^N > p_a^W$ for every age a.

The age-one states and the transition are continuous and weakly increasing in (γ, r) under this order. The same induction gives continuity and weak monotonicity of p_a^τ in (γ, r) for every age. Since $p_a^\tau \in [0,1]$ and the age weights $\lambda(1-\lambda)^a$ sum to one, dominated convergence and term-by-term monotonicity imply that

$$P_\tau(\gamma, r) = \lambda \sum_{a=0}^{\infty} (1-\lambda)^a p_a^\tau$$

is continuous and weakly increasing in (γ, r). Summing the strict age-by-age inequalities gives $P_N(\gamma, r) > P_W(\gamma, r)$. □

B.3.3 Proofs of the main text results

Proof of Proposition 5.2. Stationary employment by group satisfies

$$E^g = (1-\lambda)[(1-\delta)E^g + H^g], \qquad g \in \{W, N\}.$$

Thus

$$H^g = \bar{\delta} E^g, \qquad \bar{\delta} \equiv \frac{1-(1-\lambda)(1-\delta)}{1-\lambda}.$$

Market clearing gives $H^g = U^g f(\theta_g)$. Combining this with $E^g = L^g - U^g$ yields

$$u^g \equiv \frac{U^g}{L^g} = \frac{\bar{\delta}}{\bar{\delta} + f(\theta_g)}.$$

Since f is strictly increasing and q is strictly decreasing,

$$u^N > u^W \iff \theta_N < \theta_W \iff q(\theta_N) > q(\theta_W).$$

Finally, $H^g = \bar{\delta} E^g$ implies

$$u^N > u^W \iff \frac{H^N}{L^N} < \frac{H^W}{L^W} \iff s^H < \ell.$$

□

Proof of Proposition 5.3. Let $P \equiv V^N/\bar{V}$. For any candidate $P \in (0,1)$, Lemma B.3 uniquely determines $(\theta_N(P), \theta_W(P))$ from

$$L^N Z(\theta_N(P)) = \bar{V} P, \qquad L^W Z(\theta_W(P)) = \bar{V}(1-P).$$

These tightness levels imply

$$\gamma(P) = \frac{L^N u(\theta_N(P))}{L^N u(\theta_N(P)) + L^W u(\theta_W(P))}, \qquad r(P) = \frac{q(\theta_N(P))}{q(\theta_W(P))}.$$

The firm side then returns

$$\Psi(P;\phi) = (1-\phi)P_W(\gamma(P), r(P)) + \phi P_N(\gamma(P), r(P)).$$

Any equilibrium must satisfy $P = \Psi(P;\phi)$.

Lemma B.3 implies that $\theta_N(P)$ is strictly increasing and $\theta_W(P)$ is strictly decreasing in P. Since u and q are strictly decreasing, both $\gamma(P)$ and $r(P)$ are strictly decreasing in P. Proposition 5.1 therefore implies that $\Psi(\cdot;\phi)$ is continuous and weakly decreasing. Also,

$$(1-\rho)\gamma(P) \le \Psi(P;\phi) \le \rho + (1-\rho)\gamma(P),$$

so $\Psi(P;\phi) \in (0,1)$.

As $P \downarrow 0$, $\theta_N(P) \downarrow 0$ and $\gamma(P)$ converges to a strictly positive limit, so $\liminf_{P\downarrow 0}[\Psi(P;\phi) - P] > 0$. As $P \uparrow 1$, $\theta_W(P) \downarrow 0$ and $\gamma(P)$ remains below one, so $\limsup_{P\uparrow 1}[\Psi(P;\phi) - P] < 0$. Continuity gives existence of a fixed point. Since Ψ is weakly decreasing, $\Psi(P;\phi) - P$ is strictly decreasing, so the fixed point is unique.

Conversely, suppose $P = \Psi(P;\phi)$. Use $(\gamma(P), \theta_N(P), \theta_W(P))$ to generate the cohort paths. The fixed point equation makes aggregate targeted vacancies equal to $\bar{V}P$ and $\bar{V}(1-P)$. By construction, $V^g = U^g(P)\theta_g(P)$ and hence matching-market clearing holds. Aggregating the cohort recursion gives

$$\widetilde{H}^g = \bar{\delta}\widetilde{E}^g.$$

The matching construction also gives $\widetilde{H}^g = U^g(P)f(\theta_g(P)) = \bar{\delta}E^g(P)$, so $\widetilde{E}^g = E^g(P)$ and all accounting conditions hold. Thus fixed points and stationary equilibria are equivalent.

For the sign result, Lemma B.3 gives

$$P < \ell \iff \theta_N(P) < \theta_W(P).$$

Since u is strictly decreasing,

$$P < \ell \iff u(\theta_N(P)) > u(\theta_W(P)).$$

Proposition 5.2 gives the equivalent statement in terms of s^H. At a fixed point, these are the equilibrium unemployment and hiring shares. □

Lemma B.4 (Firm-side targeting at parity). *Suppose the aggregate targeted vacancy share is $P = \ell$. Then the implied aggregate conditions are*

$$\gamma(\ell) = \ell, \qquad r(\ell) = 1.$$

At these conditions, the firm side returns

$$\Psi(\ell;\phi) = \ell + \rho D(\delta,\lambda,\rho)(\phi - \ell),$$

where

$$D(\delta,\lambda,\rho) \equiv \lambda \sum_{a=0}^{\infty} (1-\lambda)^a M_a(\rho) \in (0,1).$$

Hence $\Psi(\ell;\phi) - \ell$ *has the same sign as* $\phi - \ell$.

Proof. At $P = \ell$,

$$\frac{\bar{V}\ell}{L^N} = \frac{\bar{V}(1-\ell)}{L^W} = \frac{\bar{V}}{L^N + L^W}.$$

The symmetric benchmark and Lemma B.3 imply $\theta_N(\ell) = \theta_W(\ell) = 1$, so $\gamma(\ell) = \ell$ and $r(\ell) = 1$.

With $r = 1$, total scaled firm size is the same for both founder types. Let $s_0 = 0$ and

$$s_{a+1} = (1-\delta)s_a + 1.$$

Define

$$\Lambda_a(\rho) \equiv \frac{(1-\delta)s_a + \rho}{(1-\delta)s_a + 1}, \qquad M_0(\rho) = 1, \qquad M_{a+1}(\rho) = \Lambda_a(\rho) M_a(\rho).$$

The composition recursion becomes

$$\pi_{a+1}^{\tau} - \ell = \Lambda_a(\rho)(\pi_a^{\tau} - \ell),$$

so, using $\pi_0^W = 0$ and $\pi_0^N = 1$,

$$\pi_a^W = \ell - \ell M_a(\rho), \qquad \pi_a^N = \ell + (1-\ell)M_a(\rho).$$

Therefore

$$p_a^W = \ell - \rho\ell M_a(\rho), \qquad p_a^N = \ell + \rho(1-\ell)M_a(\rho).$$

Averaging over founder types and firm ages gives

$$\Psi(\ell;\phi) = \lambda \sum_{a=0}^{\infty} (1-\lambda)^a [(1-\phi)p_a^W + \phi p_a^N] = \ell + \rho D(\delta,\lambda,\rho)(\phi - \ell).$$

Since $M_0(\rho) = 1$ and $0 \leq M_a(\rho) < 1$ for $a \geq 1$, $D(\delta,\lambda,\rho) \in (0,1)$. $\square$

Proof of Proposition 5.4. For fixed ϕ, write

$$F_{\phi}(P) \equiv \Psi(P;\phi) - P.$$

By Proposition 5.3, F_ϕ is strictly decreasing and has a unique zero at $P^*(\phi)$. Lemma B.4 gives

$$F_\phi(\ell) = \rho D(\delta, \lambda, \rho)(\phi - \ell).$$

Thus the unique fixed point lies below, at, or above ℓ according as $\phi < \ell$, $\phi = \ell$, or $\phi > \ell$. Proposition 5.3 then converts this comparison into the corresponding unemployment ranking, proving the threshold statement.

Now let $\phi_2 > \phi_1$. For every $P \in (0,1)$,

$$\Psi(P;\phi_2) - \Psi(P;\phi_1) = (\phi_2 - \phi_1)[P_N(\gamma(P), r(P)) - P_W(\gamma(P), r(P))] > 0,$$

where the strict inequality follows from Proposition 5.1. Evaluating at $P^*(\phi_1)$ gives $F_{\phi_2}(P^*(\phi_1)) > 0$. Since F_{ϕ_2} is strictly decreasing, its zero satisfies $P^*(\phi_2) > P^*(\phi_1)$. Hence $P^*(\phi)$ is strictly increasing in ϕ.

Finally, Lemma B.3 implies that $\theta_N(P)$ rises and $\theta_W(P)$ falls with P. Since u is strictly decreasing, $u^N(P) - u^W(P)$ is strictly decreasing in P. Because $P^*(\phi)$ is strictly increasing in ϕ, the equilibrium gap $u^N - u^W$ is strictly decreasing in ϕ. □

For the equilibrium associated with ϕ, define

$$\gamma_\phi = \gamma(P^*(\phi)), \qquad r_\phi = r(P^*(\phi)),$$

$$F_r(x) \equiv \frac{rx}{1 + (r-1)x}, \qquad p^W_{0,\phi} = (1-\rho)\gamma_\phi, \qquad p^N_{0,\phi} = (1-\rho)\gamma_\phi + \rho.$$

The average initial nonwhite share of completed hires is

$$s^H_0(\phi) \equiv (1-\phi)F_{r_\phi}(p^W_{0,\phi}) + \phi F_{r_\phi}(p^N_{0,\phi}),$$

and the Jensen comparison term is

$$\tilde{s}^H_0(\phi) \equiv F_{r_\phi}((1-\phi)p^W_{0,\phi} + \phi p^N_{0,\phi}).$$

Lemma B.5 (Founder-specific cohort paths converge to a common target)**.** *Fix founder share ϕ and let*

$$\gamma_\phi \equiv \gamma(P^*(\phi)), \qquad r_\phi \equiv \frac{q(\theta_N(P^*(\phi)))}{q(\theta_W(P^*(\phi)))}.$$

Suppose

$$\rho \max\{r_\phi, r_\phi^{-1}\} < 1.$$

Then the map

$$G_\phi(\pi) \equiv F_{r_\phi}\big((1-\rho)\gamma_\phi + \rho\pi\big)$$

has a unique fixed point $\pi^*(\phi) \in [0,1]$. For each founder type $\tau \in \{W, N\}$, the cohort composition *path generated by* (B.4)–(B.6) *satisfies*

$$\pi_a^\tau \to \pi^*(\phi) \qquad \text{as } a \to \infty.$$

Consequently, the completed nonwhite share of hires also converges to $\pi^*(\phi)$.

Proof. By Proposition B.1, the contraction condition implies that G_ϕ has a unique fixed point $\pi^*(\phi)$.

Fix founder type τ. Using (B.4)–(B.6),

$$\pi_{a+1}^\tau = (1-\omega_a^\tau)\pi_a^\tau + \omega_a^\tau G_\phi(\pi_a^\tau), \qquad \omega_a^\tau \equiv \frac{h_a^{\tau,N} + h_a^{\tau,W}}{n_{a+1}^\tau} \in (0,1].$$

Since $p_a^\tau \in [0,1]$, per-period hires satisfy

$$\bar{v}\min\{q_\phi^N, q_\phi^W\} \le h_a^{\tau,N} + h_a^{\tau,W} \le \bar{v}\max\{q_\phi^N, q_\phi^W\},$$

where $q_\phi^g \equiv q(\theta_g(P^*(\phi)))$. Because $\delta > 0$, (B.5) implies

$$n_a^\tau \le \frac{\bar{v}\max\{q_\phi^N, q_\phi^W\}}{\delta} \qquad \text{for all } a,$$

so

$$\omega_a^\tau \ge \delta \frac{\min\{q_\phi^N, q_\phi^W\}}{\max\{q_\phi^N, q_\phi^W\}} \equiv \underline{\omega}_\phi > 0.$$

Let $\kappa_\phi \equiv \rho\max\{r_\phi, r_\phi^{-1}\} < 1$. Since G_ϕ is κ_ϕ-Lipschitz and $G_\phi(\pi^*(\phi)) = \pi^*(\phi)$,

$$|\pi_{a+1}^\tau - \pi^*(\phi)| \le \big[1 - \omega_a^\tau(1-\kappa_\phi)\big]|\pi_a^\tau - \pi^*(\phi)| \le \big[1 - \underline{\omega}_\phi(1-\kappa_\phi)\big]|\pi_a^\tau - \pi^*(\phi)|.$$

The bracketed term is strictly less than one, so $\pi_a^\tau \to \pi^*(\phi)$.

Finally, the completed nonwhite share of hires is $G_\phi(\pi_a^\tau)$, so it also converges to $\pi^*(\phi)$. □

Proof of Proposition 5.5. The contraction condition implies that

$$\pi \mapsto F_{r_\phi}((1-\rho)\gamma_\phi + \rho\pi)$$

has a unique fixed point $\pi^*(\phi)$. Let

$$h_\phi(\pi) \equiv F_{r_\phi}((1-\rho)\gamma_\phi + \rho\pi) - \pi.$$

As in Proposition B.1, h_ϕ is strictly decreasing, so $h_\phi(\pi)$ has the sign of $\pi^*(\phi) - \pi$.

If $\phi = \ell$, then the threshold result gives $\gamma_\ell = \ell$ and $r_\ell = 1$. Hence $\pi^*(\ell) = \ell$ and

$$s_0^H(\ell) = \tilde{s}_0^H(\ell) = \ell.$$

Suppose $\phi < \ell$. The threshold result gives $u^N > u^W$, so Proposition 5.2 gives $r_\phi > 1$. Also,

$$\gamma_\phi - \ell = \frac{\ell(1-\ell)(u^N - u^W)}{\ell u^N + (1-\ell)u^W} > 0,$$

so $\gamma_\phi > \ell > \phi$. Since $r_\phi > 1$,

$$h_\phi(\gamma_\phi) = F_{r_\phi}(\gamma_\phi) - \gamma_\phi = \frac{\gamma_\phi(1-\gamma_\phi)(r_\phi - 1)}{1 + (r_\phi - 1)\gamma_\phi} > 0,$$

which implies $\pi^*(\phi) > \gamma_\phi$. Therefore

$$(1-\phi)p^W_{0,\phi} + \phi p^N_{0,\phi} = (1-\rho)\gamma_\phi + \rho\phi < (1-\rho)\gamma_\phi + \rho\pi^*(\phi).$$

Since F_{r_ϕ} is increasing,

$$\tilde{s}^H_0(\phi) < \pi^*(\phi).$$

Because $F''_r(x) = -2r(r-1)/[1+(r-1)x]^3$, F_{r_ϕ} is concave when $r_\phi > 1$. Jensen's inequality gives

$$s^H_0(\phi) \le \tilde{s}^H_0(\phi) < \pi^*(\phi).$$

The case $\phi > \ell$ reverses the inequalities. Then $u^N < u^W$, $r_\phi < 1$, and $\gamma_\phi < \ell < \phi$. The same argument gives $\pi^*(\phi) < \gamma_\phi$, the average initial targeted share is above $(1-\rho)\gamma_\phi + \rho\pi^*(\phi)$, and monotonicity of F_{r_ϕ} implies $\tilde{s}^H_0(\phi) > \pi^*(\phi)$. Since F_{r_ϕ} is convex when $r_\phi < 1$, Jensen's inequality gives

$$s^H_0(\phi) \ge \tilde{s}^H_0(\phi) > \pi^*(\phi).$$

Combining the three cases proves the stated sign equivalence. □

B.4 Affine extension

Sections 2 and 5 focus on the linear benchmark

$$p(\pi, \gamma) = (1-\rho)\gamma + \rho\pi.$$

This subsection states the corresponding results under the broader affine class from Appendix B.2. Throughout, we maintain $\ell \le 1/2$.

Let

$$p(\pi, \gamma) = a(\gamma) + b(\gamma)\pi = \pi p_N(\gamma) + (1-\pi)p_W(\gamma),$$

with $a(\gamma) = p_W(\gamma)$, $b(\gamma) = p_N(\gamma) - p_W(\gamma)$, and $0 \le b(\gamma) \le \bar{b} < 1$.

Proposition B.6 (Affine extension of the firm-level results)**.** *Fix γ and $q_N, q_W > 0$. If*

$$b(\gamma)\frac{\max\{q_N, q_W\}}{\min\{q_N, q_W\}} < 1,$$

then $\widehat{p}(\cdot, \gamma; q_N, q_W)$ has a unique fixed point. If $\pi_0^N > \pi_0^W$, the four predictions in section 2 continue to hold. When $q_N = q_W$,

$$\mathbb{E}[\pi_{k+1} \mid \pi_k, \gamma] = \pi_k + \frac{[1 - b(\gamma)][\bar{\pi}(\gamma) - \pi_k]}{k+1}, \qquad \bar{\pi}(\gamma) = \frac{a(\gamma)}{1 - b(\gamma)}.$$

The affine rule preserves founder effects and convergence. Relative to the linear benchmark, only two objects change: the common target is $\bar{\pi}(\gamma)$ rather than γ, and the persistence term is $b(\gamma)$ rather than a constant ρ.

Proof. The contraction and coupling arguments from Propositions B.1 and B.2 apply with $b(\gamma)$ in place of ρ. □

Proposition B.7 (Affine extension of the macro fixed point results)**.** *Suppose the affine model satisfies the same contraction condition as the benchmark, with $\bar{b}$ in place of ρ:*

$$\bar{b} \sup_{P \in (0,1)} \max\{r(P), r(P)^{-1}\} < 1.$$

Then Proposition 5.1, Proposition 5.2, and Proposition 5.3 continue to hold. In particular, stationary equilibrium is again characterized by a unique fixed point $P^*(\phi)$, the map $\phi \mapsto P^*(\phi)$ is strictly increasing, and

$$u^N > u^W \iff P^*(\phi) < \ell \iff s^H < \ell.$$

The parity comparison differs. In the linear benchmark, founder parity implies unemployment parity: $P^*(\ell) = \ell$. In the affine case this identity generally fails. If $\bar{\pi}(\ell) \leq \ell$, then $P^*(\ell) \leq \ell$. Equivalently, founder parity is no longer enough to guarantee unemployment parity, and any equilibrium with $u^N \leq u^W$ must have $\phi \geq \ell$.

Proof. The proof of Proposition 5.1 only uses that $p(\pi, \gamma)$ is continuous, weakly increasing in γ, and affine with nonnegative slope in π. The proofs of Propositions 5.2 and 5.3 then apply without change.

For the parity comparison, evaluate the firm side at $P = \ell$. Then $\gamma = \ell$ and $r = 1$. Because total hires per period are the same for both founder types when $r = 1$, the founder-weighted mean composition

$$\bar{\pi}_a(\phi) \equiv (1 - \phi)\pi_a^W + \phi\pi_a^N$$

follows the same linear recursion as a single cohort with initial condition $\bar{\pi}_0(\phi) = \phi$ and target $\bar{\pi}(\ell)$. If $\phi = \ell$ and $\bar{\pi}(\ell) \leq \ell$, then $\bar{\pi}_a(\ell) \leq \ell$ for every age. The corresponding founder-weighted targeted share is therefore

$$a(\ell) + b(\ell)\bar{\pi}_a(\ell) \leq a(\ell) + b(\ell)\ell \leq \ell,$$

where the last inequality is equivalent to $\bar{\pi}(\ell) \leq \ell$. Averaging over firm ages gives $\Psi(\ell; \ell) \leq \ell$, so the fixed point also satisfies $P^*(\ell) \leq \ell$. Since $P^*(\phi)$ is strictly increasing in ϕ, any equilibrium with $P^*(\phi) \geq \ell$ must have $\phi \geq \ell$. □

For the equilibrium associated with founder share ϕ, define

$$\gamma_\phi \equiv \gamma(P^*(\phi)), \qquad r_\phi \equiv \frac{q(\theta_N(P^*(\phi)))}{q(\theta_W(P^*(\phi)))}, \qquad F_r(x) \equiv \frac{rx}{1+(r-1)x}.$$

Let $\pi^*(\phi)$ denote the unique fixed point of

$$\pi \mapsto F_{r_\phi}(a(\gamma_\phi) + b(\gamma_\phi)\pi),$$

and let the average initial nonwhite share of completed hires be

$$s_0^H(\phi) \equiv (1-\phi)F_{r_\phi}(a(\gamma_\phi)) + \phi F_{r_\phi}(a(\gamma_\phi) + b(\gamma_\phi)).$$

Corollary B.8 (Life-cycle profiles in the affine case)**.** *Maintain the hypotheses of Proposition B.7 and suppose $\bar{\pi}(\gamma) \leq \gamma$ for every $\gamma \in [0, \ell]$. Then*

$$\pi^*(\phi) > s_0^H(\phi) \qquad \Longrightarrow \qquad u^N > u^W.$$

Thus, when nonwhites make up at most half of the labor force, a rising average life-cycle profile still implies that co-racial hiring raises nonwhite unemployment. The converse need not hold.

Proof. Suppose instead that $u^N \leq u^W$. Then Proposition B.7 gives $P^*(\phi) \geq \ell$, so $\gamma_\phi \leq \ell$, $r_\phi \leq 1$, and $\phi \geq \ell$. By assumption,

$$\bar{\pi}(\gamma_\phi) \leq \gamma_\phi \leq \ell \leq \phi.$$

When $r_\phi \leq 1$, the map F_{r_ϕ} is increasing, convex, and satisfies $F_{r_\phi}(x) \leq x$. Therefore Jensen's inequality gives

$$s_0^H(\phi) \geq F_{r_\phi}(a(\gamma_\phi) + b(\gamma_\phi)\phi) \geq F_{r_\phi}(\bar{\pi}(\gamma_\phi)).$$

Also,

$$\pi^*(\phi) = F_{r_\phi}(a(\gamma_\phi) + b(\gamma_\phi)\pi^*(\phi)) \leq F_{r_\phi}(\bar{\pi}(\gamma_\phi)),$$

because $F_{r_\phi}(x) \leq x$ implies $\pi^*(\phi) \leq \bar{\pi}(\gamma_\phi)$. Hence $s_0^H(\phi) \geq \pi^*(\phi)$, so the average life-cycle profile cannot be rising. Taking the contrapositive yields the result. □

Relative to the linear benchmark, the firm-level convergence results and the scalar fixed point equilibrium survive under the broader affine class. What is special to the benchmark is the exact identity

$$\operatorname{sign}(\pi^*(\phi) - s_0^H(\phi)) = \operatorname{sign}(u^N - u^W) = \operatorname{sign}(\ell - \phi).$$

In the affine case, the common target becomes $\bar{\pi}(\gamma)$ and founder parity need not imply unemployment parity. Under the same-group mechanisms above and $\ell \leq 1/2$, however, the one-sided implication used in the paper remains valid: a rising average life-cycle profile implies $u^N > u^W$.

C Additional exhibits

C.1 Estimating slope of relationship between nonwhite share and cumulative hires

We estimate a more parametric variant of equation (3) 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

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

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.6 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 (3), which describes the relationship between NONWHITE_{jh} and cumulative hires for all firms pooled 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 4.1 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.3.

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.4.

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:

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

The η^n coefficient estimates are plotted in Panel B of Figure C.4.

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

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

Table C.2
Characteristics of hires at entrant firms

		White Founder			Nonwhite Founder		
Hires:	All	All	White	Nonwhite	All	White	Nonwhite
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
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 3.2. Columns 2–4 limit to entrant firms with white founders and columns 5–7 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.3
Characteristics of hires at entrant firms, balanced panel

		White Founder			Nonwhite Founder		
Hires:	All	All	White	Nonwhite	All	White	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 3.2. 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 2–4 limit to entrant firms with white founders and columns 5–7 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 (1)	< 50% Nonwhite Incumbents (%) (2)	≥ 50% Nonwhite Incumbents (%) (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

Hires:	All (1)	< 50% Nonwhite Incumbents (%) All (4)	< 50% Nonwhite Incumbents (%) White (5)	< 50% Nonwhite Incumbents (%) Nonwhite (6)	≥ 50% Nonwhite Incumbents (%) All (7)	≥ 50% Nonwhite Incumbents (%) White (8)	≥ 50% Nonwhite Incumbents (%) Nonwhite (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. Columns 2–4 limit to new establishment subsidiaries where less than 50% of incumbent firm employees are nonwhite and columns 5–7 limit to new establishment subsidiaries where 50% or more of incumbent firm employees are nonwhite. 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

SLOPE ESTIMATES FOR RELATIONSHIP BETWEEN NONWHITE SHARE AND 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	0.0392
× White Founder					(0.0003)	(0.0003)	(0.0003)	(0.0004)
log Cumulative Hires (N)					-0.0157	-0.01888	-0.0180	-0.0154
× 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							

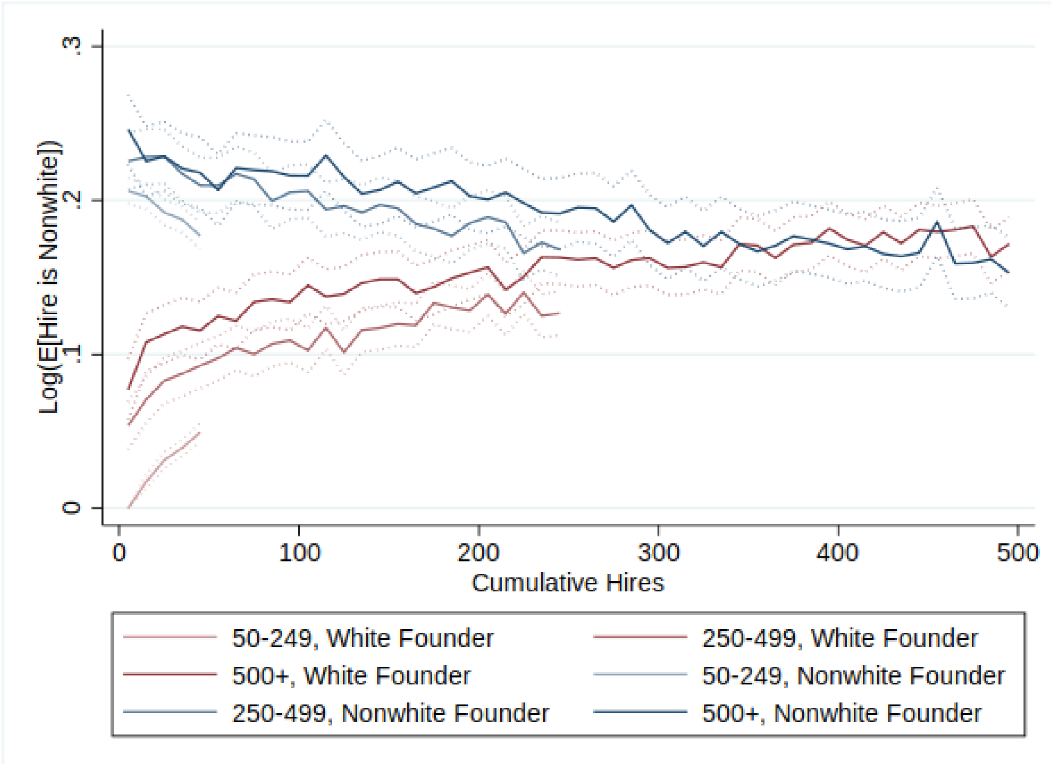
Note: This table presents δ coefficient estimates from equation (C.1), which summarizes the relationship between the racial composition of a firm's hires and its cumulative hires to date. The outcome is an indicator for whether a hire is nonwhite. All specifications include year and firm fixed effects. 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.7
Characteristics of hires at new subsidiary establishments, balanced panel

		< 50% Nonwhite Incumbents (%)			≥ 50% Nonwhite Incumbents (%)		
Hires:	All	All	White	Nonwhite	All	White	Nonwhite
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
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

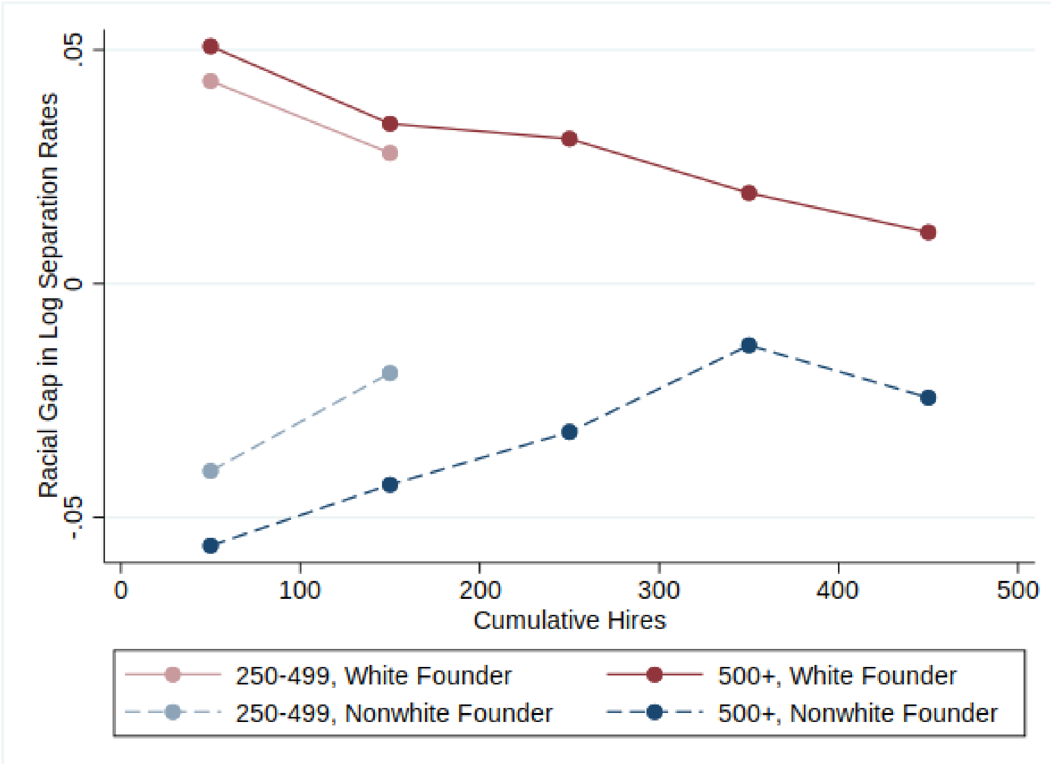
This table reports summary statistics for hires at a balanced panel of new subsidiary establishments in *Relação Anual de Informações Sociais* (RAIS) data for the years 2003–2017. We limit the sample to private sector, indeterminate-length contracts. Each observation is a worker-firm job spell. We restrict 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 2–4 limit to new establishment subsidiaries where less than 50% of incumbent firm employees are nonwhite and columns 5–7 limit to new establishment subsidiaries where 50% or more of incumbent firm employees are nonwhite. 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.

Figure C.1
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.2
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 hires, 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.3

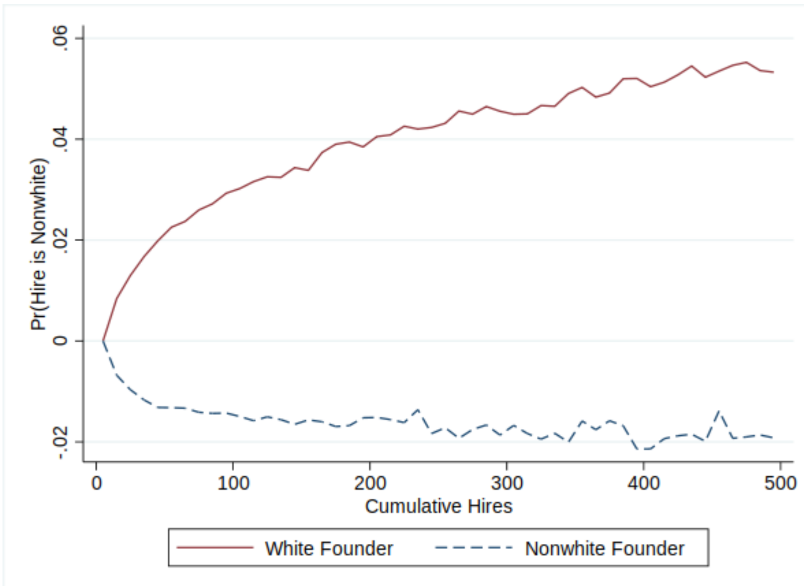
Nonwhite share of hires increases over life cycle, linear probability model

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 replace 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.

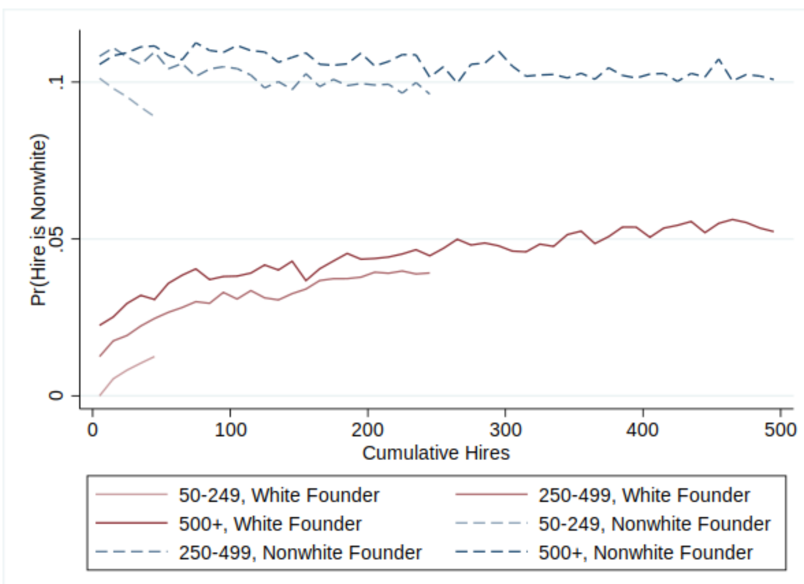
Figure C.4

Founder race and convergence in nonwhite share of hires, linear probability model

(a) All entrants

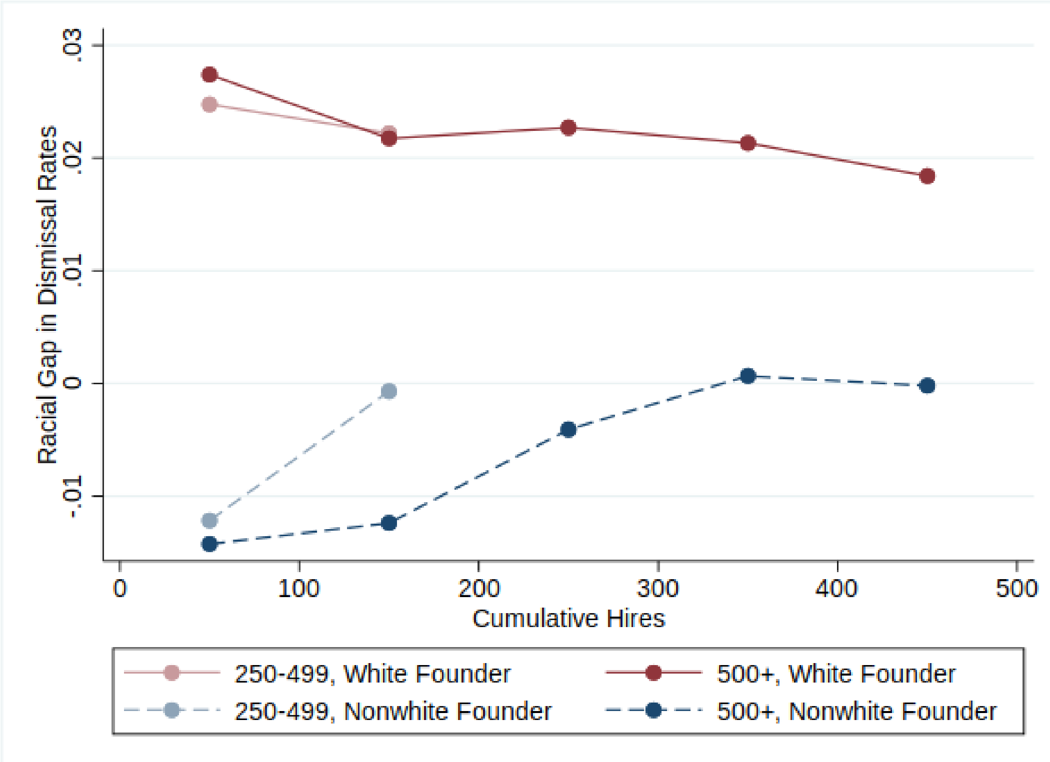


(b) Balanced panel of entrants



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. In both panels we exclude the inferred founder from the new hires we consider and when measuring cumulative hires. In both panels 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.5
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). 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.6

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 (9), where the model is estimated separately for pairs of worker and founder race.

D Evidence from U.S. aggregate firm cohorts

This appendix describes the U.S. aggregate cohort exercise summarized in the introduction and discussion. We use public Quarterly Workforce Indicators (QWI) tabulations from the U.S. Census Bureau derived from Longitudinal Employer-Household Dynamics (LEHD) data. The outcome we use is `HirN`, the count of new hires. We download quarterly, seasonally unadjusted tabulations by race, firm-age bin, broad NAICS sector, and geography for 2003–2017. The numerator is the count for workers classified as Black or African American alone, and the denominator is the count for all races. The state analysis includes all states and the District of Columbia. The county analysis requests the 100 counties with the largest Census 2000 populations.

The public firm-age bins do not form a panel of firms. We therefore construct aggregate firm birth cohorts from the age bins that can be aligned over time. We first aggregate quarterly Black and total new hires to annual geography-by-industry-by-firm-age cells. Let a denote the observed firm-age bin and let $\mathcal{E}_{t,a}$ denote the implied two-year entry interval for firms observed in calendar year t at age a. We assign

$$\mathcal{E}_{t,a} = \begin{cases} [t-1, t] & \text{if } a = 0\text{–}1, \\ [t-3, t-2] & \text{if } a = 2\text{–}3, \\ [t-5, t-4] & \text{if } a = 4\text{–}5. \end{cases}$$

An aggregate firm cohort is a geography-by-industry-by-entry-interval cell. For example, the age-0–1 cell observed in year t, the age-2–3 cell observed in year $t+2$, and the age-4–5 cell observed in year $t+4$ are assigned to the same aggregate cohort. We exclude the 6–10 and 11-plus firm-age bins because their wider intervals would mix multiple birth cohorts. We keep source quarterly cells for which the Census status flag indicates valid, unsuppressed hire counts for both the numerator and denominator and for which total new hires are at least 500. This minimum-denominator rule removes cells that are ex ante most likely to be suppressed. After aggregating the remaining quarterly cells to annual cohort-by-age cells, we restrict all descriptive and regression results to balanced aggregate cohorts observed in all three age bins.

Table D.8 summarizes the resulting profiles. Panel A reports hires-weighted Black shares of new hires by firm-age bin. In the state tabulations, the Black share of new hires rises from 13.82 percent at ages 0–1 to 16.69 percent at ages 4–5. In the large-county tabulations, it rises from 15.24 percent to 18.95 percent. Panel B follows the same balanced aggregate cohorts across age bins. From ages 0–1 to ages 4–5, the Black share of new hires rises by 1.19 percentage points in the state sample and by 1.30 percentage points in the county sample.

We next estimate linear and Poisson specifications on the same balanced aggregate cohort sample. Let H_{ca}^{Black} denote Black new hires in aggregate cohort c and firm-age bin a, let H_{ca} denote all new hires, and let $S_{ca} = H_{ca}^{\text{Black}}/H_{ca}$ denote the Black share of new hires. The linear specification is

$$S_{ca} = \sum_{b \in \{2\text{–}3, 4\text{–}5\}} \beta^b \times \mathbb{1}_{\{a=b\}} + \lambda_c + \varepsilon_{ca}, \tag{D.1}$$

TABLE D.8
BLACK SHARE OF NEW HIRES IN U.S. QWI
AGGREGATE FIRM AGE COHORTS

	States	Top 100 counties
Panel A. Hires-weighted Black share of new hires		
0–1 years	13.82	15.24
2–3 years	15.89	17.61
4–5 years	16.69	18.95
Panel B. Matched aggregate cohort changes		
0–1 to 2–3 years	0.99	0.99
0–1 to 4–5 years	1.19	1.30
2–3 to 4–5 years	0.26	0.38
Annual cells	12,354	8,418
Aggregate cohorts	4,118	2,806
Matched transition cells	12,354	8,418

Note: Panel A reports hires-weighted Black shares of new hires, in percent. Panel B reports changes in percentage points for matched geography-industry aggregate birth cohorts, weighted by baseline total new hires. The data are public Census Quarterly Workforce Indicators tabulations for 2003–2017. State tabulations include the District of Columbia. County tabulations request the 100 counties with the largest Census 2000 populations. All rows use source quarterly cells with total new hires of at least 500 and balanced aggregate cohorts observed in all three displayed age bins.

where λ_c is an aggregate cohort fixed effect. The omitted category is ages 0–1. We weight equation (D.1) by total new hires in the cell. As a proportional specification, we estimate the conditional Poisson model

$$\log(E(H_{ca}^{\text{Black}} \mid \cdot)) = \log H_{ca} + \sum_{b \in \{2\text{–}3, 4\text{–}5\}} \theta^b \times \mathbb{1}_{\{a=b\}} + \lambda_c. \tag{D.2}$$

This specification treats $\log H_{ca}$ as the exposure offset and estimates within-cohort proportional differences in the Black share of new hires.

Table D.9 reports the estimates. In the linear specification, the Black share of new hires is about 1.2 percentage points higher at ages 2–3 than at ages 0–1, and about 1.7 to 1.9 percentage points higher at ages 4–5. The conditional Poisson estimates are positive in both geographies, again indicating a rising Black share of new hires as aggregate cohorts age.

TABLE D.9
U.S. QWI AGGREGATE COHORT REGRESSIONS FOR BLACK NEW HIRES

	Linear probability model		Poisson	
	States (1)	Counties (2)	States (3)	Counties (4)
2–3 years	1.177	1.237	0.0803	0.0767
	(0.007)	(0.012)	(0.0028)	(0.0038)
4–5 years	1.671	1.879	0.1099	0.1106
	(0.007)	(0.013)	(0.0038)	(0.0048)
Aggregate cohort FEs	Yes	Yes	Yes	Yes
N Observations	12,354	8,418	12,354	8,418
N Aggregate cohorts	4,118	2,806	4,118	2,806
Total hires	153,167,457	47,410,778	153,167,457	47,410,778

Note: Columns 1 and 2 report estimates from equation (D.1), in percentage points. Columns 3 and 4 report the Poisson log coefficients $\hat{\theta}$ from equation (D.2). Standard errors are in parentheses. The omitted age bin is 0–1 years. Heteroskedasticity-robust standard errors are reported. All columns use source quarterly cells with total new hires of at least 500 and balanced aggregate cohorts observed in all three displayed age bins.

These estimates should be interpreted as suggestive evidence rather than a U.S. replication of the main research design. The QWI tabulations do not identify specific firms nor do they measure cumulative hires within firms. Changes over time within aggregate cohorts may reflect within-firm hiring dynamics, differential survival, differential growth, or other changes in the composition of firms within geography-industry cells.