

Discrimination in Hiring: Evidence from Retail Sales

Online Appendix

Alan Benson, Simon Board, and Moritz Meyer-ter-Vehn

B Supplementary Discussion and Tables

B.1 Sensitivity

In this section we discuss the sensitivity of our results in several dimensions.

Productivity. In Section 4.2 we estimate our baseline productivity regression (11) via weighted-least squares. Tables 3 and 4 shows the resulting supermodularity statistics and confidence intervals from 10,000 bootstraps. As stated in the main text, mean productivity is supermodular, especially for white-Hispanic and black-Hispanic pairs, and productivity variance is log-submodular, especially for black-Hispanic pairs. Thus, the black-Hispanic pairs indicate screening discrimination (97%) while white-Hispanic indicate screening or complementarity (99.1%). Table B1 shows the regression coefficients underlying these results. As discussed in Section 4.2, Table B2 conducts several robustness checks and Figure B2 recalculates Figure 4 for (a) the first six months of workers’ tenure, and (b) later months.

Turnover. Section 4.3 considered turnover after 6 months, showing that $\Delta_{\Theta\Theta'}^x < 0$ for all three pairs, with white-Hispanic being significant. The underlying regression is shown in Table B3. Table B4(a) shows that the results are robust to measuring turnover at 3, 9 and 12 months. All the coefficients are negative except black-Hispanic at 6 months; only the white-Hispanic coefficients are significantly negative.

Table B4(b) then includes the worker’s average productivity, $\bar{e}_i$, as a control. As expected, this has a negative sign; e.g. with the “6 month” regression, the coefficient is -0.0609 , meaning a one standard deviation increase in $\bar{e}_i$ of 0.439 lowers turnover by 2.7% from a base of 36.1%. In the model, a worker is fired when their realized productivity falls below a threshold. Under screening discrimination and complementary production, same-race workers have higher productivity than cross-race workers, so controlling for SPH should increase $\hat{\Delta}_{\Theta\Theta}^x$ towards zero. Under taste-based discrimination, same-race workers have lower productivity than cross-race workers, so controlling for SPH should lower $\hat{\Delta}_{\Theta\Theta}^x$. Table B4(b) shows that $\hat{\Delta}_{\Theta\Theta}^x$ tends to increase for white-Hispanic and black-Hispanic pairs.

Full sample of workers. Our baseline results in Section 4 focus on 48,755 WBH newly hired salespeople for whom we have productivity data. We can also run the hiring regression (9) and 6-month turnover regression (15) on the entire sample of 812,399 WBH newly hired workers that includes cashiers, backroom operators, and service technicians. Table B5 shows our hiring results. As before, managers are significantly more likely to hire same-race workers. Specifically, the gap between same-race and cross-race hiring $\hat{\beta}_{\Theta\theta}^r - \hat{\beta}_{\Theta\theta}^c$ is 2.5pp, 3.2pp and 2.4pp for white, black and Hispanic workers, which is similar to the corresponding gaps for salespeople (3.9pp, 2.9pp and 4.0pp). Table B6 shows turnover is submodular for each pair of races; these results are significant at the 1% level. The size of these coefficients is $(\hat{\Delta}_{WB}^x + \hat{\Delta}_{WH}^x + \hat{\Delta}_{BH}^x)/6 = -2.9\%$ compared with -0.8% for salespeople. Table B4(c) shows these results are robust measuring at turnover at 3, 9 and 12 months.

Internal transfers. Our test of discrimination in Table 4 uses the productivity of a worker under their hiring manager. If workers were reassigned randomly (e.g. Glover, Pallais, and Pariente, 2017), then we could use productivity under non-hiring managers to test for complementarity. Unfortunately, this is not the case. First, managers employ about 4.7pp more same-race candidates than one would expect from random matching (albeit with large standard errors). Second, the race of the hiring manager and the new manager are correlated (after controlling for store and department fixed effects). Our random assignment assumption (see Section 2.5) may be satisfied after controlling for the race of a worker’s hiring manager (and possibly their productivity under the hiring manager). One could then apply our discrimination tests to internal transfers by comparing the supermodularity of hiring and productivity conditional on the hiring manager’s race. Unfortunately we do not have enough transfers to tightly estimate such parameters. In addition, transfers might depend on workers’ preferences, which complicates the theory.

B.2 Alternative theories

This paper has considered three theories (taste-based discrimination, screening discrimination, complementary production). Here we briefly discuss other factors that might be at play.

Customer bias. Our main theories suppose that worker productivity is determined by the hiring decisions of their managers (and directly by the manager, in the case of complementarity). If manager race is correlated with the race of customers, biased customer preferences may look like complementary production between the manager and worker. To be specific, if there are “white towns” with predominantly white customers and white managers, and “Hispanic towns” with predominantly Hispanic customers and Hispanic managers, then same-race preferences among customers will induce $\hat{\Delta}_{WH}^y > 0$ even without complementarity between managers and workers.³⁷ To address this,

³⁷We allow for store fixed effects, which accounts for the fact that all workers may be more productive in some

we re-run our productivity regression (11) and include the fraction of the population of the same race as the worker, which serves as a proxy for the race of customers. Column (2) of Table B8 shows the “share same-race pop” coefficient is positive (1.5pp) but not significant. Including this variable has little impact on the supermodularity statistics, $\hat{\Delta}_{\Theta\theta}^y$ or $\hat{\Delta}_{\Theta\theta}^v$, suggesting that customer bias is not driving the results in Tables 3-4.

Peer effects. Analogous to customer bias, if manager race is correlated with the race of other team members, complementarity between peers may look like complementarity between the manager and worker. To be specific, if white managers are in charge of predominantly white teams and Hispanic managers are in charge of predominantly Hispanic teams, then complementarity between peers will induce $\hat{\Delta}_{WH}^y > 0$ even without complementarity between managers and workers. To address this, we re-run our productivity regression (11) and include the fraction of peers of the same race as the worker. Column (3) of Table B8 shows the “share same-race worker peers” coefficient is positive (1.8pp) and significant at the 10% level. However, including this variable does not qualitatively change the supermodularity statistics ($\hat{\Delta}_{\Theta\theta}^y, \hat{\Delta}_{\Theta\theta}^v$), suggesting that peer effects are not driving the results in Tables 3-4. With this said, the interpretation of this regression is unclear since the race of a worker’s peers is endogenous and the positive “worker peers” coefficient may pick up managerial heterogeneity. For example, if a manager has strong same-race signals, the average worker will have more same-race peers and the workers will have high productivity.³⁸

Favoritism. Our theory allows for complementarity, whereby a worker is more productive with a same-race manager. This is different from favoritism, whereby a manager gives preferential treatment to same race workers. The latter is intrinsically zero-sum, e.g. assigning same-race workers to the best shifts. Our measure of productivity controls for the worker’s shift via their target sales, eliminating the most obvious source of favoritism. In order to test for residual favoritism, we note that if the manager gives the best roles to same-race workers then any given worker’s productivity should improve if there are fewer peers who share the same race as the manager. So inspired, we re-run our productivity regression (11) and include the fraction of peers of the same race as the manager. Column (4) of Table B8 shows the “share same-race manager peers” coefficient is positive (2.6pp) and significant at the 5% level; in contrast, the favoritism hypothesis would predict a negative coefficient. This is evidence against favoritism but, as with the peer effects, this result may

stores than others. However, in the white/Hispanic town example, the stores might have equal average productivity.

³⁸One may also wonder if the race of peers are correlated with managers’ hiring decisions. Table B7 includes peer race in our baseline hiring regression and shows that both managers’ race and peer race is correlated with the race of new hires. Again, the interpretation of this is unclear. Having more peers of one race may bias the manager towards future hires of the same race: For example, Benson and LePage (2022) use the current data set to examine how a manager’s hiring decision depends on their past hires. Alternatively, since the mix of peers is endogenous, the positive peer race coefficient may pick up heterogeneity in managerial bias: If a manager is very biased, their teams and new hires will both be more likely to be same-race.

come from reverse causation since the race of a worker’s peers is endogenous and the positive “manager peers” coefficient may pick up managerial heterogeneity. For example, if a manager has strong same-race signals, she will have more same-race peers and her workers will have high productivity.

Referral networks. As discussed in Section 2.5, our model assumes that each manager selects from an equal distribution of applicants $(n_{\Theta}, \mu_{\Theta}, \sigma_{\Theta}^2)$ after controlling for store, department and month fixed effects. This excludes the possibility that the manager may obtain referrals through their social networks. If referrals have the same productivity distribution $(\mu_{\Theta}, \sigma_{\Theta}^2)$ as other applicants of that race, the larger number of applicants n_{Θ} would raise the number of same-race hires (Table 2) but would not affect productivity (Tables 3-4). However, referrals may also be positively selected (e.g. Beaman and Magruder, 2012; Pallais and Sands, 2016). If the manager receives valuable information from their social contacts, this would lower σ_{Θ}^2 and look like our screening model. Alternatively, if referrals have a higher mean distribution μ_{Θ} , this would look like our complementary production model.

It is hard to evaluate the importance of referrals since we do not observe applications. At an anecdotal level, one HR representative told us that, in her view, it is rare for managers to know applicants. We can also look for indirect evidence of referrals. Brown, Setren, and Topa (2016) find that “most referrals take place between a provider and a recipient with similar characteristics in terms of age, gender, ethnicity, education, and division and staff level within the corporations.” Thus, if more of the same-race workers are referrals, one might expect the same-race workers share other characteristics with the manager. Table B9 looks at the average age and gender differences between the worker and manager by each racial pair; there is no evidence that same-race workers are more similar to their managers.

Worker preferences. The model of taste-discrimination assumes that managers are biased towards same-race workers. An alternative model is that managers are neutral, but workers are biased towards same-race managers, accepting the offer with a higher probability when the manager shares the same race. Such a model also predicts the same managers hire more same-race candidates. The impact on productivity depends on the details of the model. If workers’ decision to reject offers are independent of their skill, the resulting productivity is modular, $\Delta_{\Theta\theta}^y = 0$. Alternatively, if an applicant’s propensity to accept a job offer rises in their productivity (because of the commission) and in the race-match with their recruiting manager (because of taste-based discrimination), the logic underlying Proposition 1 implies that productivity is submodular. Ultimately, we cannot rule out such forces, but they seem less important with hiring decisions than retention or transfer decisions.

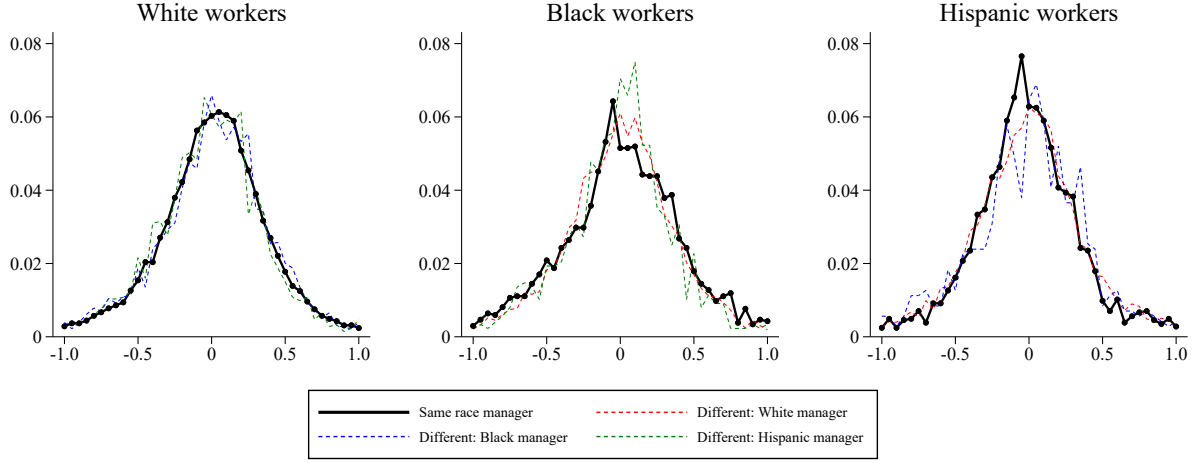


FIGURE B1: HISTOGRAMS OF WORKER PRODUCTIVITY BY MANAGER AND WORKER RACE
Notes: Recall that $\bar{e}_i$ is the within-worker mean residual from the productivity regression (11). By construction, $\bar{e}_i$ has zero mean within each race-pair $\Theta\theta$, so $\bar{\varepsilon}_i := \bar{e}_i + \hat{\beta}_{\Theta\theta}^y$ is the productivity of worker i , after controlling for store, department, month and tenure. The figure plots the histogram of $\bar{\varepsilon}_i$ across manager and worker race. For example, the left hand figure shows the distribution of productivity for white workers; the black line is productivity under white managers, the green line is productivity under black managers, and the blue line productivity under Hispanic managers. We plot the frequency distributions for every combination of racial pair at 5%-wide intervals of productivity.

TABLE B1: **BASILINE PRODUCTIVITY REGRESSION**

Worker-Manager race	Prod. Mean, $\hat{\beta}_{\Theta\theta}^y$	90% Conf. Int.	Prod. Variance, $\hat{\sigma}_{\Theta\theta}^2$	No. workers
White-White	reference		0.114	26,847
White-Black	-0.0019	[-0.0172, 0.0155]	0.120	2,285
White-Hispanic	-0.0260	[-0.0415, -0.0073]	0.111	2,122
Black-White	-0.0251	[-0.0358, -0.0139]	0.143	6,172
Black-Black	-0.0130	[-0.0297, 0.0050]	0.143	2,344
Black-Hispanic	-0.0323	[-0.0572, -0.0045]	0.129	877
Hispanic-White	-0.0163	[-0.0283, -0.0050]	0.119	4,557
Hispanic-Black	-0.0383	[-0.0695, -0.0076]	0.172	710
Hispanic-Hispanic	-0.0146	[-0.0281, 0.0013]	0.105	2,841
Productivity Mean:				
White-Black ($\hat{\Delta}_{WB}^y$)	0.0140	[-0.0117, 0.0383]		37,648
White-Hispanic ($\hat{\Delta}_{WH}^y$)	0.0277	[0.0049, 0.0510]		36,367
Black-Hispanic ($\hat{\Delta}_{BH}^y$)	0.0429	[-0.0011, 0.0880]		6,772
Productivity Variance:				
White-Black ($\hat{\Delta}_{WB}^v$)	0.953	[0.799, 1.131]		37,648
White-Hispanic ($\hat{\Delta}_{WH}^v$)	0.914	[0.759, 1.110]		36,367
Black-Hispanic ($\hat{\Delta}_{BH}^v$)	0.679	[0.500, 0.956]		6,772
Location FEs	Yes			
Month FEs	Yes			
Department FEs	Yes			
Observations	336,046			48,755

Notes: This table shows the outcome of the baseline productivity regression (11) that underlies Tables 3-4, estimated via WLS, inversely weighted to tenure. The top section shows the regression coefficients $\hat{\beta}_{\Theta\theta}^y$ by race-pair and the corresponding cross-agent variance $\hat{\sigma}_{\Theta\theta}^2$ given by (13). The second section shows the mean supermodularity coefficient $\hat{\Delta}_{\Theta\Theta}^y$, as defined by (12), while the third section shows the variance supermodularity coefficient $\hat{\Delta}_{\Theta\Theta}^v$, as defined by (14). The confidence intervals come from 10,000 bootstraps.

TABLE B2: DISCRIMINATION TESTS: SENSITIVITY

Statistic	WLS-T (1)	WLS-NT (2)	WLS-3T (3)	WLS-6T (4)	OLS-T (5)	OLS-NT (6)	OLS-3T (7)	OLS-6T (8)
Productivity Mean:								
White-Black ($\hat{\Delta}_{WB}^y$)	0.0140 [-0.0117, 0.0383]	0.0143 [-0.0126, 0.0382]	0.0181 [-0.0033, 0.0402]	0.0126 [-0.0072, 0.0326]	-0.0025 [-0.0274, 0.0237]	-0.0025 [-0.0264, 0.0231]	0.0039 [-0.0208, 0.0300]	-0.0012 [-0.0242, 0.0217]
White-Hispanic ($\hat{\Delta}_{WH}^y$)	0.0277 [0.0049, 0.0510]	0.0301 [0.0076, 0.0529]	0.0253 [0.0049, 0.0450]	0.0212 [0.0024, 0.0393]	-0.0012 [-0.0240, 0.0236]	0.0003 [-0.0227, 0.0254]	0.0125 [-0.0109, 0.0360]	0.0105 [-0.0111, 0.0313]
Black-Hispanic ($\hat{\Delta}_{BH}^y$)	0.0429 [-0.0011, 0.0880]	0.0454 [-0.0001, 0.0863]	0.0283 [-0.0083, 0.0626]	0.0267 [-0.0103, 0.0641]	0.0163 [-0.0244, 0.0557]	0.0179 [-0.0217, 0.0598]	0.0278 [-0.0151, 0.0720]	0.0264 [-0.0159, 0.0688]
Productivity Variance:								
White-Black ($\hat{\Delta}_{WB}^v$)	0.953 [0.799, 1.131]	0.950 [0.797, 1.148]	0.921 [0.772, 1.102]	1.005 [0.833, 1.185]	0.929 [0.780, 1.099]	0.927 [0.775, 1.099]	0.915 [0.761, 1.095]	0.996 [0.846, 1.189]
White-Hispanic ($\hat{\Delta}_{WH}^v$)	0.914 [0.759, 1.110]	0.915 [0.765, 1.131]	0.931 [0.772, 1.138]	0.999 [0.822, 1.199]	0.911 [0.759, 1.088]	0.909 [0.740, 1.089]	0.933 [0.765, 1.118]	1.004 [0.827, 1.223]
Black-Hispanic ($\hat{\Delta}_{BH}^v$)	0.679 [0.500, 0.956]	0.680 [0.503, 0.941]	0.840 [0.620, 1.195]	0.798 [0.592, 1.103]	0.664 [0.493, 0.933]	0.665 [0.490, 0.939]	0.843 [0.616, 1.209]	0.800 [0.574, 1.111]
White-Black joint test (%):								
$\hat{\Delta}^y > 0, \hat{\Delta}^v < 1$ (screen)	57.9	58.7	73.5	45.1	30.9	34.4	45.6	24.8
$\hat{\Delta}^y > 0, \hat{\Delta}^v > 1$ (comple)	23.6	23.1	18.3	41.1	10.3	8.7	13.2	21.8
$\hat{\Delta}^y < 0, \hat{\Delta}^v < 1$ (s & t)	10.7	9.6	5.3	4.5	45.8	43.4	31.6	24.0
$\hat{\Delta}^y < 0, \hat{\Delta}^v > 1$ (taste)	7.8	8.6	2.9	9.3	13.0	13.5	9.6	29.4
White-Hispanic joint test (%):								
$\hat{\Delta}^y > 0, \hat{\Delta}^v < 1$ (screen)	75.6	78.3	71.2	50.7	36.0	40.3	57.7	36.1
$\hat{\Delta}^y > 0, \hat{\Delta}^v > 1$ (comple)	22.1	20.7	27.6	46.0	12.3	9.7	24.1	42.4
$\hat{\Delta}^y < 0, \hat{\Delta}^v < 1$ (s & t)	1.4	0.8	0.7	1.7	41.6	40.4	12.8	11.0
$\hat{\Delta}^y < 0, \hat{\Delta}^v > 1$ (taste)	0.9	0.2	0.5	1.6	10.1	9.6	5.4	10.5
Black-Hispanic joint test (%):								
$\hat{\Delta}^y > 0, \hat{\Delta}^v < 1$ (screen)	91.8	92.4	73.5	78.5	72.7	77.1	67.0	75.8
$\hat{\Delta}^y > 0, \hat{\Delta}^v > 1$ (comple)	2.6	2.5	17.4	9.8	1.6	1.6	18.3	10.2
$\hat{\Delta}^y < 0, \hat{\Delta}^v < 1$ (s & t)	5.2	4.6	6.2	9.1	25.1	20.6	9.9	10.5
$\hat{\Delta}^y < 0, \hat{\Delta}^v > 1$ (taste)	0.4	0.5	2.9	2.6	0.6	0.7	4.8	3.5

Notes: This table reports the supermodularity statistics (as in Table 3) and discrimination tests (as in Table 4) for seven model variants. Column (1) is the baseline specification, using weighted-least squares and a tenure control. Column (2) dispenses with the tenure control while Columns (3) and (4) use only the first three and six months of each workers tenure, respectively. Columns (5)-(8) show the corresponding results for ordinary-least-squares. Below each coefficient is the 90% confidence interval computed from 1000 bootstraps (except for Column (1), which uses 10,000 bootstraps).

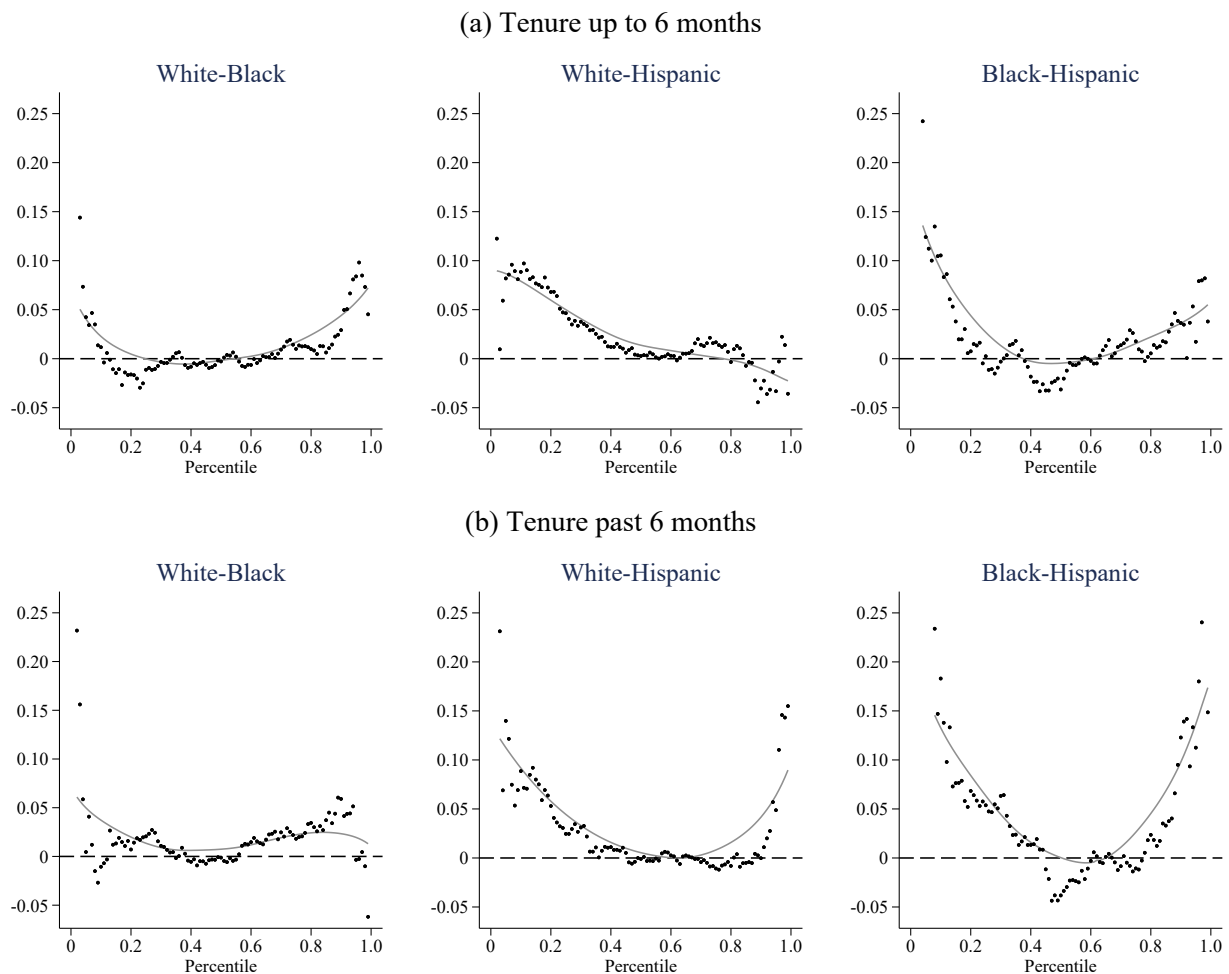


FIGURE B2: **SUPERMODULARITY OF PRODUCTIVITY BY QUANTILE**

Notes: This figure recreates Figure 4 with data above/below six months. Specifically, we run the baseline productivity regression (11) and take the residuals e_{it} . We then truncate the data, calculate i 's average residual $\bar{e}_i$ and calculate $\hat{\Delta}_{\Theta\Theta}^{y^q}$, as defined in (19). This figure shows across 98 one-percentile bins, after we have thrown away the top and bottom bins. The Lowess curve has bandwidth 0.8.

TABLE B3: **BASELINE TURNOVER REGRESSION**

6 month turnover	Point estimate	90% confidence interval	Number of workers
White-White ($\hat{\beta}_{Ww}^x$)	reference		21,613
White-Black ($\hat{\beta}_{Wb}^x$)	0.0033	[-0.0155, 0.0221]	1,885
White-Hispanic ($\hat{\beta}_{Wh}^x$)	-0.0128	[-0.0326, 0.0069]	1,694
Black-White ($\hat{\beta}_{Bw}^x$)	0.0425	[0.0302, 0.0548]	5,007
Black-Black ($\hat{\beta}_{Bb}^x$)	0.0304	[0.0116, 0.0491]	1,930
Black-Hispanic ($\hat{\beta}_{Bh}^x$)	-0.0157	[-0.0455, 0.0141]	717
Hispanic-White ($\hat{\beta}_{Hw}^x$)	-0.0073	[-0.0213, 0.0067]	3,699
Hispanic-Black ($\hat{\beta}_{Hb}^x$)	-0.0104	[-0.0436, 0.0228]	572
Hispanic-Hispanic ($\hat{\beta}_{Hh}^x$)	-0.0555	[-0.0733, -0.0378]	2,222
Location FEs	Yes		
Month FEs	Yes		
Department FEs	Yes		
Observations	39,339		
Mean turnover	0.361		
Pairwise relations			
White-Black ($\hat{\Delta}_{WB}^x$)	-0.0154	[-0.0436, 0.0127]	30,435
White-Hispanic ($\hat{\Delta}_{WH}^x$)	-0.0354	[-0.0698, -0.0011]	29,228
Black-Hispanic ($\hat{\Delta}_{BH}^x$)	0.0009	[-0.0589, 0.0607]	5,441

Notes: This table shows the baseline turnover regression (15) that underlies the turnover results in Table 3. The supermodularity coefficient $\hat{\Delta}_{WB}^x$ is defined by (16).

TABLE B4: **TURNOVER SUPERMODULARITY STATISTICS: SENSITIVITY**

	White-Black ($\hat{\Delta}_{WB}^x$)		White-Hispanic ($\hat{\Delta}_{WH}^x$)		Black-Hispanic ($\hat{\Delta}_{BH}^x$)	
	Point est.	90% conf. int.	Point est.	90% conf. int.	Point est.	90% Conf. Int.
(a) Salespeople						
3 months	-0.0144	[-0.0391, 0.0103]	-0.0130	[-0.0434, 0.0174]	-0.0316	[-0.0841, 0.0209]
6 months	-0.0154	[-0.0436, 0.0127]	-0.0354	[-0.0698, -0.0011]	0.0009	[-0.0589, 0.0607]
9 months	-0.0221	[-0.0520, 0.0077]	-0.0444	[-0.0806, -0.0082]	-0.0293	[-0.0924, 0.0338]
12 months	-0.0224	[-0.0529, 0.0082]	-0.0442	[-0.0813, -0.0071]	-0.0417	[-0.1063, 0.0229]
(b) Salespeople, controlling for productivity						
3 months	-0.0270	[-0.0533, -0.0006]	-0.0004	[-0.0338, 0.0330]	-0.0315	[-0.0874, 0.0244]
6 months	-0.0206	[-0.0507, 0.0094]	-0.0053	[-0.0431, 0.0324]	0.0176	[-0.0460, 0.0813]
9 months	-0.0341	[-0.0660, -0.0021]	-0.0099	[-0.0499, 0.0301]	-0.0136	[-0.0811, 0.0540]
12 months	-0.0284	[-0.0612, 0.0045]	-0.0087	[-0.0498, 0.0324]	-0.0149	[-0.0844, 0.0546]
(c) All workers						
3 months	-0.0289	[-0.0353, -0.0225]	-0.0566	[-0.0645, -0.0487]	-0.0491	[-0.0621, -0.0361]
6 months	-0.0382	[-0.0458, -0.0307]	-0.0730	[-0.0823, -0.0636]	-0.0641	[-0.0795, -0.0487]
9 months	-0.0379	[-0.0454, -0.0304]	-0.0887	[-0.0979, -0.0795]	-0.0789	[-0.0941, -0.0637]
12 months	-0.0314	[-0.0385, -0.0243]	-0.0945	[-0.1033, -0.0857]	-0.0800	[-0.0944, -0.0656]

Notes: This table shows how the supermodularity coefficient for the turnover regression (15) depends on our assumptions. Part (a) shows the regression for salespeople where $TURN_i$ is classified as leaving after 3, 6, 9 and 12 months. The “6 month” row corresponds to the baseline regression in Table B3. Part (b) shows the same regressions for salespeople, where we add the person’s average productivity $\bar{e}_i$ as a control variable, as defined in Section 3.1. The coefficient on $\bar{e}_i$ is stable across the regressions; for example, in the “6 month” model the coefficient on $\bar{e}_i$ is -0.0609 with a standard error of 0.0060 . Part (c) shows the same regression as part (a), for all workers rather than just salespeople.

TABLE B5: **HIRING TESTS FOR ALL WORKERS**

	White worker	Black worker	Hispanic worker
White manager	57.5 [57.4, 57.6]	25.2 [25.1, 25.3]	17.3 [17.2, 17.4]
Black manager	54.5 [54.2, 54.9]	28.0 [27.7, 28.3]	17.5 [17.2, 17.7]
Hispanic manager	55.6 [55.3, 56.0]	24.5 [24.2, 24.8]	19.8 [19.6, 20.1]

Notes: This table shows the hiring regression (9) for 812,399 newly hired WBH workers hired by WBH managers, controlling for location, department, and month. Below each coefficient is the 90% confidence interval.

TABLE B6: **TURNOVER REGRESSION FOR ALL WORKERS**

6 month turnover	Point estimate	90% confidence interval	Number of workers
White-White ($\hat{\beta}_{Ww}^x$)	reference		331,665
White-Black ($\hat{\beta}_{Wb}^x$)	0.0548	[0.0521, 0.0575]	119,984
White-Hispanic ($\hat{\beta}_{Wh}^x$)	-0.0045	[-0.0079, -0.0011]	66,017
Black-White ($\hat{\beta}_{Bw}^x$)	0.0008	[-0.0047, 0.0064]	22,250
Black-Black ($\hat{\beta}_{Bb}^x$)	0.0174	[0.0125, 0.0222]	29,996
Black-Hispanic ($\hat{\beta}_{Bh}^x$)	-0.0378	[-0.0468, -0.0292]	8,490
Hispanic-White ($\hat{\beta}_{Hw}^x$)	-0.0100	[-0.0159, -0.0041]	19,489
Hispanic-Black ($\hat{\beta}_{Hb}^x$)	0.0320	[0.0247, 0.0394]	12,364
Hispanic-Hispanic ($\hat{\beta}_{Hh}^x$)	-0.0874	[-0.0918, -0.0831]	37,693
Location FEs	Yes		
Month FEs	Yes		
Department FEs	Yes		
Observations	648,036		
Mean turnover	0.489		
Pairwise relations			
White-Black ($\hat{\Delta}_{WB}^x$)	-0.0382	[-0.0458, -0.0307]	503,895
White-Hispanic ($\hat{\Delta}_{WH}^x$)	-0.0730	[-0.0808, -0.0651]	454,864
Black-Hispanic, ($\hat{\Delta}_{BH}^x$)	-0.0641	[-0.0770, -0.0512]	88,543

Notes: This table shows the turnover regression (15) for all newly hired WBH workers hired by WBH managers.

TABLE B7: **HIRING TESTS WITH PEER RACE**

	(1) White worker	(2) Black worker	(3) Hispanic worker
White manager	ref	-2.5 [-3.6, -1.4]	-4.0. [-5.0, -3.0]
Black manager	-2.8 [-4.1, -1.5]	ref.	-3.8 [-5.1, -2.5]
Hispanic manager	-4.3 [-5.6, -2.9]	-2.3 [-3.7, -0.9]	ref.
White peers	4.7 [-1.3, 10.6]	-2.7 [-7.7, 2.4]	-2.0 [-6.6, 2.6]
Black peers	-1.7 [-8.4, 4.9]	5.4 [-0.3, 11.0]	-3.6 [-8.7, 1.4]
Hispanic peers	1.5 [-5.3, 8.4].	-1.5 [-7.3, 4.3]	0.0 [-5.3, 5.2]
Constant	62.5 [56.9, 68.0]	22.5 [17.8, 27.3]	21.5 [17.2, 25.9]

Notes: This table expands the hiring regression (9) to include the share of white/black/Hispanic peers. As in the baseline hiring regression, we allow for store, department and month fixed effects. There are 43,164 salespeople; this is fewer than in Table 2 since we only observe team composition for people hired during the sample period. Below each coefficient is the 90% confidence interval.

TABLE B8: **DISCRIMINATION TESTS: RACE OF CUSTOMERS AND PEERS**

Statistic	Baseline (1)	Customers (2)	Worker Peers (3)	Manager Peers (4)
Productivity Mean:				
White-Black ($\hat{\Delta}_{WB}^y$)	0.0140 [-0.0117, 0.0383]	0.0112 [-0.0130, 0.0352]	0.0057 [-0.0186, 0.0319]	0.0028 [-0.0220, 0.0284]
White-Hispanic ($\hat{\Delta}_{WH}^y$)	0.0277 [0.0049, 0.0510]	0.0232 [-0.0009, 0.0460]	0.0179 [-0.0063, 0.0435]	0.0093 [-0.0140, 0.0349]
Black-Hispanic ($\hat{\Delta}_{BH}^y$)	0.0429 [-0.0011, 0.0880]	0.0396 [-0.0034, 0.0863]	0.0338 [-0.0096, 0.0785]	0.0304 [-0.0122, 0.0727]
Productivity Variance:				
White-Black ($\hat{\Delta}_{WB}^v$)	0.953 [0.799, 1.131]	0.951 [0.806, 1.120]	0.948 [0.797, 1.120]	0.923 [0.771, 1.097]
White-Hispanic ($\hat{\Delta}_{WH}^v$)	0.914 [0.759, 1.110]	0.804 [0.662, 0.978]	0.910 [0.749, 1.098]	0.933 [0.770, 1.122]
Black-Hispanic ($\hat{\Delta}_{BH}^v$)	0.679 [0.500, 0.956]	0.594 [0.435, 0.830]	0.682 [0.496, 0.954]	0.709 [0.516, 1.003]
White-Black joint test (%):				
$\Delta^y > 0, \Delta^v < 1$ (screen)	57.9	56.0	49.7	49.0
$\Delta^y > 0, \Delta^v > 1$ (comple)	23.6	21.0	17.2	10.2
$\Delta^y < 0, \Delta^v < 1$ (s & t)	10.7	14.3	19.3	26.8
$\Delta^y < 0, \Delta^v > 1$ (taste)	7.8	8.7	13.8	14.0
White-Hispanic joint test (%):				
$\Delta^y > 0, \Delta^v < 1$ (screen)	75.6	91.4	70.7	55.7
$\Delta^y > 0, \Delta^v > 1$ (comple)	22.1	3.0	17.5	20.0
$\Delta^y < 0, \Delta^v < 1$ (s & t)	1.4	5.5	9.7	16.7
$\Delta^y < 0, \Delta^v > 1$ (taste)	0.9	0.1	2.1	7.6
Black-Hispanic joint test (%):				
$\Delta^y > 0, \Delta^v < 1$ (screen)	91.8	92.7	88.3	83.9
$\Delta^y > 0, \Delta^v > 1$ (comple)	2.6	0.6	2.7	4.1
$\Delta^y < 0, \Delta^v < 1$ (s & t)	5.2	6.5	8.7	10.9
$\Delta^y < 0, \Delta^v > 1$ (taste)	0.4	0.2	0.3	1.1
Share Same-Race Population				
		0.0145 [-0.0103, 0.0425]		
Share Same-Race Worker Peers				
			0.0184 [0.0026, 0.0334]	
Share Same-Race Manager Peers				
				0.0255 [0.0111 0.0397]

Notes: This table reports the supermodularity statistics (as in Table 3) and discrimination tests (as in Table 4) for three model variants. Column (1) repeats the baseline results. Column (2) proxies for the race of customers by controlling for the share of the local population who are the same race as the worker that year. We use MSA population, or non-MSA state population if the location isn't in an MSA, from the annual American Community Survey. Column (3) controls for the share of the worker's peers who share the same race as the worker that month (taking into account non-WBH races). Column (4) controls for the share of worker's peers who share the same race as the manager that month (taking into account non-WBH races). We present 90% confidence intervals calculated via 1,000 bootstraps.

TABLE B9: **REFERRALS TEST: AGE AND GENDER DIFFERENCES BY RACIAL PAIR**

	Age difference in years			Same gender indicator		
	White worker	Black worker	Hispanic worker	White worker	Black worker	Hispanic worker
White manager	17.20 (11.30) [17.09, 17.32]	17.54 (11.50) [17.29, 17.78]	18.42 (11.51) [18.13, 18.70]	0.383 (0.486) [0.378, 0.388]	0.381 (0.486) [0.371, 0.391]	0.386 (0.487) [0.373, 0.398]
Black manager	16.78 (10.77) [16.40, 17.15]	17.05 (11.17) [16.66, 17.43]	17.33 (10.16) [16.69, 17.97]	0.382 (0.486) [0.365, 0.399]	0.411 (0.492) [0.394, 0.428]	0.392 (0.489) [0.361, 0.423]
Hispanic manager	14.82 (10.24) [14.45, 15.19]	13.75 (9.44) [13.22, 14.29]	14.66 (9.15) [14.38, 14.95]	0.386 (0.487) [0.369, 0.404]	0.411 (0.492) [0.383, 0.439]	0.426 (0.495) [0.411, 0.442]
	Cross-race mean		17.12	Cross-race mean		0.385
	Same-race mean		16.96	Same-race mean		0.389
	p-value, same means		0.160	p-value, same means		0.454

Notes: The table reports differences in ages and gender by racial pair, for salespeople in the main sample, given the race of the worker and manager. The **left panel** reports the means and standard deviation for the absolute value of difference in ages in years, along with the 90% confidence interval. For instance, the top left cell restricts the sample to white workers hired by white managers, then reports the mean absolute value of their age differences. The “cross-race mean” reports the average of the “cross-race” cells, weighted by sample size, while the “same-race” mean reports the average of the same-race cells. The p-values correspond to two-sample t-test for equal means, where the dependent variable (the absolute value of the difference in ages) depends on whether the manager is same-race or cross-race. The **right panel** reports the mean and standard deviation of an indicator that equals one if the manager and worker have different genders. A majority of employees are men, so the probability that sexes differ is less than 0.5. Data in this table include 47,094 newly hired sales workers included in the main sample, and exclude approximately 3.5% of workers for whom manager gender or age are missing.

C Omitted Proofs

C.1 Maximizing expected sales

In this appendix we analyze the model variant in which managers maximize expected sales, as discussed in Section 2.5. Let Y be a worker's sales, and assume a manager wishes to hire workers with $Y \geq Y^*$. As in the main text, let $\hat{y} = E[y|\tilde{y}]$ be expected log-sales, with mean μ , estimator variance $\eta^2 = \sigma_0^4/(\sigma_0^2 + \sigma_\epsilon^2)$, and estimation error $y - \hat{y}$ with residual variance $\gamma^2 := \sigma_0^2\sigma_\epsilon^2/(\sigma_0^2 + \sigma_\epsilon^2)$.

The manager's expected utility from a worker given estimate $\hat{y}$ is thus

$$E[Y|\hat{y}] = E[e^y|\hat{y}] = e^{\hat{y}} E[e^{y-\hat{y}}|\hat{y}] = e^{\hat{y}+\gamma^2/2}.$$

The manager thus wishes to hire a workers with estimate $\hat{y} \geq y^* := \log Y^* - \gamma^2/2$. Intuitively, managers lower their standard, $y^* < \log Y^*$, since they are risk-loving with respect to uncertainty in log-productivity.

Taste-based discrimination. In this model, biased managers gain extra utility from same-race candidates. Suppose they gain utility Y/D from a cross-race candidate, and Y from a same-race candidate, where $D > 1$ is the same-race bias. Taking logs, the cross-race cutoff $y_c^* = \log Y^* + d - \gamma^2/2$ thus exceeds the same-race cutoff $y_s^* = \log Y^* - \gamma^2/2$ by the bias $d = \log D$. The analysis of hiring is thus identical to that in Section 2.1, and thus the proofs of Propositions 1(a)-(c) carry over unchanged. When it comes to turnover, suppose that the manager fires the worker if their realized productivity y drops more than $\tilde{\tau}$ below the cutoff y_i^* . Thus the turnover rate for $i \in \{c, s\}$ is

$$x_i := \Pr(y \leq y_i^* - \tilde{\tau} | \hat{y} \geq y_i^*) = \Pr(y \leq y_i^* - \tau | \hat{y} \geq y_i^*),$$

where $\tau := \tilde{\tau} - \gamma^2/2$. This is the same as in Section 2.1, and Proposition 1(d) immediately applies.

Complementary production. As with taste-based discrimination, the analysis of the baseline case (where managers maximize expected log-sales) is essentially unchanged. Since the residual variance is fixed at γ^2 , the cross-race cutoff y_c^* exceeds the same-race cutoff y_s^* by k , and the firing threshold is τ below the hiring threshold. Proposition 3 immediately generalizes.

Screening discrimination. In this model, the signal variance σ_ϵ^2 is lower for same-race applicants than for cross-race applicants, $\sigma_s^2 \leq \sigma_c^2$. This implies that cross-race hires have more residual variance than same-race hires, $\gamma_s^2 \leq \gamma_c^2$. Thus, managers set a lower threshold for cross-race hires, $y_c^* = \log Y^* - \gamma_c^2/2 < \log Y^* - \gamma_s^2/2 = y_s^*$. Intuitively, managers are risk-loving with respect to the log-productivity, and thus likes cross-race candidates because of their high residual variance. We

first verify that managers still hire more same-race workers.

Proposition 2(a)'. Assume $\log Y^* > \mu + \sigma_0^2$. In the screening model, the hiring probability is higher for same-race applicants than cross-race applicants, $p_s > p_c$.

Proof. For part (a), let Φ and ϕ be the cdf and pdf of the standard normal distribution. Write the precision of the prior as $h_0 := 1/\sigma_0^2$ and the precision of the raw signal as $h_\epsilon = 1/\sigma_\epsilon^2$. The estimator variance and residual variance are then $\eta^2 = h_\epsilon/((h_0 + h_\epsilon)h_0)$ and $\gamma^2 = 1/(h_0 + h_\epsilon)$. Now, the acceptance rate can be written as a function of the signal precision

$$\begin{aligned} p(h_\epsilon) &:= \Pr\left(\hat{y} \geq \log Y^* - \frac{\gamma^2}{2}\right) = 1 - \Phi\left(\frac{\log Y^* - \mu - \frac{\gamma^2}{2}}{\eta}\right) = \\ &= 1 - \Phi\left(\sqrt{\frac{(h_0 + h_\epsilon)h_0}{h_\epsilon}}(\log Y^* - \mu) - \frac{1}{2}\sqrt{\frac{h_0}{(h_0 + h_\epsilon)h_\epsilon}}\right) \end{aligned}$$

Simple computations show that the derivative equals

$$\begin{aligned} p'(h_\epsilon) &= -\phi(\cdot) \times \sqrt{h_0} \left(-\frac{h_0}{2h_\epsilon} \sqrt{\frac{1}{h_\epsilon(h_0 + h_\epsilon)}} (\log Y^* - \mu) + \frac{1}{4} \left(\frac{1}{h_\epsilon} + \frac{1}{h_0 + h_\epsilon} \right) \sqrt{\frac{1}{(h_0 + h_\epsilon)h_\epsilon}} \right) \\ &= \phi(\cdot) \times \frac{1}{2} \sqrt{\frac{h_0}{(h_0 + h_\epsilon)h_\epsilon}} \frac{h_0}{h_\epsilon} \left((\log Y^* - \mu) - \frac{1}{2} \frac{1}{h_0} \left(1 + \frac{h_\epsilon}{h_0 + h_\epsilon} \right) \right) \end{aligned}$$

Given that $h_\epsilon/(h_0 + h_\epsilon) \leq 1$, this last term is positive if $\log Y^* \geq \mu + \sigma_0^2$. Since same-race candidates have greater signal precision h_ϵ , they have a higher acceptance rate, and so managers hire relatively more same-race workers. $\square$

Turning to productivity, the proof of Proposition 2(b),(c) shows that productivity of cross-race hires has lower mean and greater variance when using the same bar, $y_c^* = y_s^*$. When risk-neutral managers lower the bar for cross-race recruits, this further lowers their expected productivity and raises productivity variance. Finally, when we consider turnover, the proof of Proposition 2(d) shows that cross-race hires have higher turnover when using the same bar, $y_c^* = y_s^*$. When we lower the bar for cross-race recruits, this further raises their turnover.

C.2 Proof of Proposition 6

Let p_s and p_c be the probability same-race and cross-race candidates are hired. Let $M_t \in [0, 1]$ and $N_t \in [0, 1]$ be the number of A managers and workers at a firm, and consider the dynamics (N_t, M_t) in (N, M) space (see Figure C1).

First, consider the dynamics of workers, given the the number of managers. Let ψ be the rate at which workers and managers leave; this could be due to exit or promotion. A race- A manager

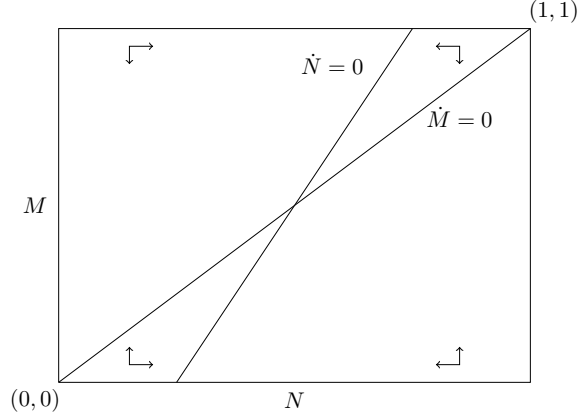


FIGURE C1: **POPULATION DYNAMICS OF MANAGERS AND WORKERS**

hires race-A workers over race-B workers in the ratio $p_s n / p_c(1 - n)$. Hence the number of race-A workers evolves according to

$$\dot{N} = \psi \left[(1 - M) \left(\frac{p_c n}{p_c n + p_s(1 - n)} \right) + M \left(\frac{p_s n}{p_s n + p_c(1 - n)} \right) - N \right]. \quad (31)$$

Thus, in (N, M) space, $\dot{N} = 0$ is a straight line from $(\frac{p_c n}{p_c n + p_s(1 - n)}, 0)$ to $(\frac{p_s n}{p_s n + p_c(1 - n)}, 1)$, so with slope greater than one.

Next, consider the dynamics of managers, given the number of workers. Managers retire at rate ψ and replacements are chosen randomly from the population of workers, so $\dot{M} = \psi(N - M)$. Thus, $\dot{M} = 0$ along the diagonal where $N = M$.

Figure C1 illustrates the dynamics of (N, M) . It is evident that there is a unique steady state where $\dot{M} = 0$ and $\dot{N} = 0$, and that the system converges to this steady state from any initial condition. When $N = M = n$, then

$$\dot{N} = \psi \left[(1 - n) \left(\frac{p_c n}{p_c n + p_s(1 - n)} \right) - n \left(\frac{p_c(1 - n)}{p_s n + p_c(1 - n)} \right) \right] \stackrel{\text{sgn}}{=} (p_s - p_c)(2n - 1) > 0$$

if $n > 1/2$. Thus the steady state must satisfy $N^* = M^* > n$.

Finally, we consider how many managers of a particular race are required to obtain equitable hiring. Rearranging (31), if we have $\dot{N} = 0$ when $N = n$, then M must solve

$$M \left(\frac{(p_s^2 - p_c^2)n(1 - n)}{(p_s n + p_c(1 - n))(p_c n + p_s(1 - n))} \right) = \frac{(p_s - p_c)(1 - n)n}{p_c n + p_s(1 - n)}.$$

Letting $1 - n = \epsilon \approx 0$, we obtain

$$M \left(\frac{p_s^2 - p_c^2}{p_s p_c} \right) \epsilon \approx \frac{p_s - p_c}{p_c} \epsilon,$$

so that the number of race-B managers is $1 - M \approx p_c / (p_s + p_c)$.

References

- BEAMAN, L., AND J. MAGRUDER (2012): “Who gets the Job Referral? Evidence from a Social Networks Experiment,” *American Economic Review*, 102(7), 3574–3593.
- BENSON, A., AND L.-P. LEPAGE (2022): “Learning to Discriminate on the Job,” Working paper, University of Minnesota.
- BROWN, M., E. SETREN, AND G. TOPA (2016): “Do Informal Referrals Lead to Better Matches? Evidence from a Firm’s Employee Referral System,” *Journal of Labor Economics*, 34(1), 161–209.
- GLOVER, D., A. PALLAIS, AND W. PARIENTE (2017): “Discrimination as a Self-fulfilling Prophecy: Evidence from French Grocery Stores,” *Quarterly Journal of Economics*, 132(3), 1219–1260.
- PALLAIS, A., AND E. G. SANDS (2016): “Why the Referential Treatment? Evidence from Field Experiments on Referrals,” *Journal of Political Economy*, 124(6), 1793–1828.