Social Capital and Racial Inequality

Nicholas Tenev*

November 9, 2017

Abstract

This paper estimates the extent to which inequality in jobs found through friends can account for the aggregate wage gap between black workers and others in the US. Data from the NLSY79 are used to estimate a job search model in which individual productivity is distinguished from social capital by comparing the wages and frequency of jobs found directly with those of jobs found through friends. Jobs found through friends tend to pay more, but this premium is lower for black workers; the difference can account for more than a tenth of the racial wage gap.

*University of Wisconsin–Madison. This paper has benefitted enormously from discussions with Naoki Aizawa, Steven Durlauf, Chao Fu, Rasmus Lentz, Corina Mommaerts, Chris Taber, Matt Wiswall, and also from comments from seminar participants at the University of Wisconsin–Madison. Most recent draft available at https://sites.google.com/site/nicholastenev/tenev_JMP.pdf

1 Introduction

Social connections play an important role in finding work.[1] But as not everyone has the same friends, some people may end up in higher-paying jobs than others, even if their qualifications are otherwise identical. This paper's contribution is to empirically assess the importance of inequality in social capital to aggregate wage inequality between groups. In particular, I focus on the oft-studied[2] persistent inequality in labor market outcomes between black workers and others in the US. I estimate that at least a tenth of the wage gap between these groups can be explained by differences in jobs found through friends, even after controlling for individual productivity.

The model is simple and similar to that of Burdett (1978), but with heterogeneous workers, wage growth with experience, and two types of offers. Workers receive job offers through direct search and also through friends. The chance of receiving an offer depends on its source, as well as worker characteristics including current employment status. A worker will accept an offer if it maximizes expected discounted utility from earnings. At the end of each period workers may lose their jobs; the chance of this happening depends on worker characteristics as well as how the job was found.

The parameters are estimated jointly via method of simulated moments on panel data from the NLSY79, which in certain years asks respondents whether their jobs were found through friends. This is the key to the identification strategy: the wages

[1] Granovetter (1973) found that roughly half of jobs are found through a social connection; more recent work on the importance of referrals includes Ioannides and Datcher Loury (2004) and Schmutte (2016).

[2] See Altonji and Blank (1999) for a review.

of jobs found directly provide a measure of individuals' productivity, allowing the relative wages of jobs found through friends to be used as a measure of social capital. I find that the job offers black workers find through friends pay less than those of equally productive non-black workers.

Montgomery (1991; 1992) showed how persistent inequality in wages and educational attainment can arise between two groups even if they have equal productive potential. More recently, Calvó-Armengol and Jackson (2004) explored a similar idea in an explicit network setting. The key to these papers is that if two groups are more likely to form in-group social connections (homophily), their labor market outcomes can follow different trajectories. Workers choose higher education or labor force participation because their friends are doing the same, and the payoffs to education or participation are higher if you have friends who can help you find a job (strategic complementarity). In a similar vein, Arrow and Borzekowski (2004) argue that plausible differences in network degree (number of connections) may account for black-white income disparity. I take these theoretical foundations to the data, and find that differences in jobs found through friends are indeed an important part of aggregate racial inequality in labor market outcomes.

This paper remains agnostic on exactly why friends are helpful in job search. The results are consistent with a model like that of Calvó-Armengol and Jackson (2004), in which friends are useful for merely alerting you that an opening exists, but also consistent with Simon and Warner (1992), in which referrals are valuable because they provide information to the hiring firm.[3] As my data cannot distin-

[3]Brown et al. (2015) find that panel data from a large US corporation support a model such as that of in which referrals are valuable because they convey information about workers' productivity. For example, jobs found through referrals tend to pay higher initial wages–a finding corroborated by this paper.

guish between these models, firm behavior is not modeled: arrival rates and wage distributions are reduced-form specifications flexible enough to qualitatively match a variety of possible explanations for why friends matter on the job market. The focus is instead on measuring how much the fruits of search through friends differ by race.

My results are broadly consistent with previous empirical work on the topic. Schmutte (2015) uses geographic variation to show that social interactions can explain why some workers get higher paying jobs. Holzer (1987) finds that differences in job finding rates can account for most of the racial difference in unemployment outcomes for youths, and Green et al. (1999) find that jobs found through friends pay less for black workers in certain cities. This paper makes two main contributions relative to previous empirical work. The first is scope—I use nationally representative data and consider both unemployment and wages of early- and mid-career workers. Second, I allow the wage premium of jobs found through friends to vary with human capital, to avoid misinterpreting group differences in human capital as differences in social capital (see Section 4.2 for further discussion).

Prior accounts of racial inequality in labor market outcomes have often focused on differences in skills, such as formal education or hidden investment in human capital that both causes and is caused by statistical discrimination (such as Coate and Loury (1993)[4]). This paper is qualitatively different in that it studies inequality in labor market outcomes controlling for individual productivity, so the policy implications may be quite different. For example, affirmative action in hiring may be able to balance out social capital inequality and provide disadvantaged groups with

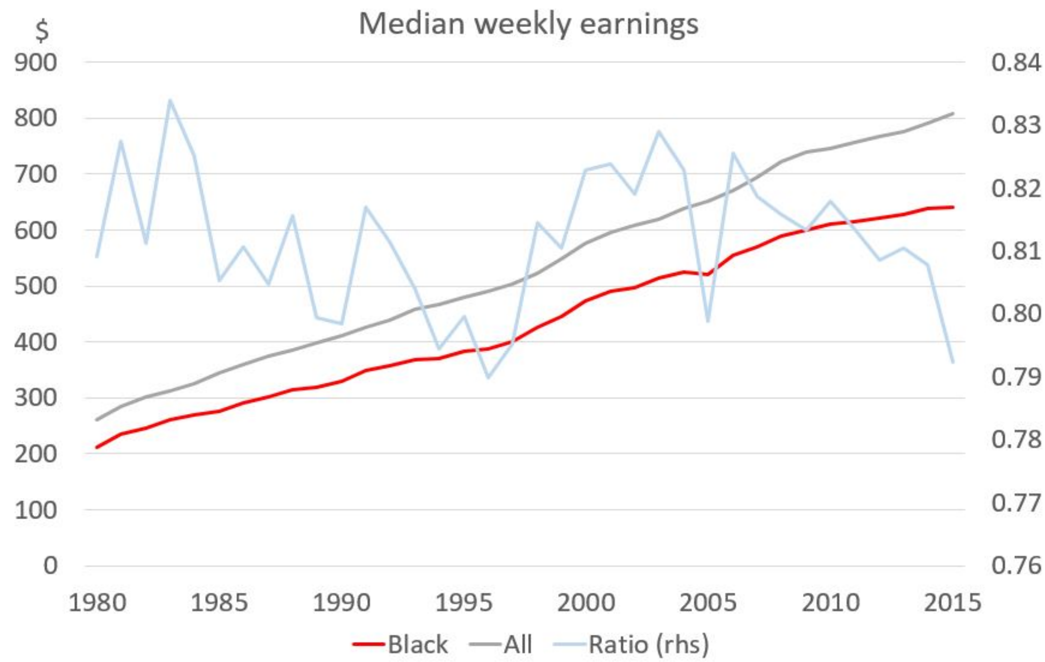
[4] Fang and Moro (2011) review other work in this vein.

the same rewards to education enjoyed by groups with better social connections. Furthermore, it may increase the social capital of the disadvantaged group: the more minority workers have high-status jobs, the better the job-finding prospects of their friends. This can explain empirical evidence from Miller (2017) that firms which increase their hiring of black workers to meet affirmative action standards required by government contracts tend to continue to do so even after they are no longer compelled to do so. Finally, while racial inequality is the focus of this paper, the importance of social capital may prove important for understanding inequality in other dimensions as well.

Other work in this area provides complementary explanations for persistent inequality. Becker and Tomes (1979) and Loury (1981) show that inequality can persist if lower-income parents invest less in their children's education. Durlauf (1996) shows that gaps between neighborhoods can persist where education is a local public good. This paper establishes the importance of network job-finding alongside these other mechanisms as a key driver of racial inequality. Whereas most previous work focuses on inequality driven by differences in human capital, this paper isolates differences in job-finding which may owe entirely to differences in social connections unrelated to human capital or productivity.

Section 2 motivates the research, reviewing stylized facts about race, labor market outcomes, and social connections. Section 3 develops the labor market model. Section 4 describes the data used and performs preliminary analysis that roughly encapsulates the main result in a single fixed-effects regression, for those short on time. Section 5 explains the estimation and identification strategy, and Section 6 presents the main results.

Figure 1: Racial inequality in earnings



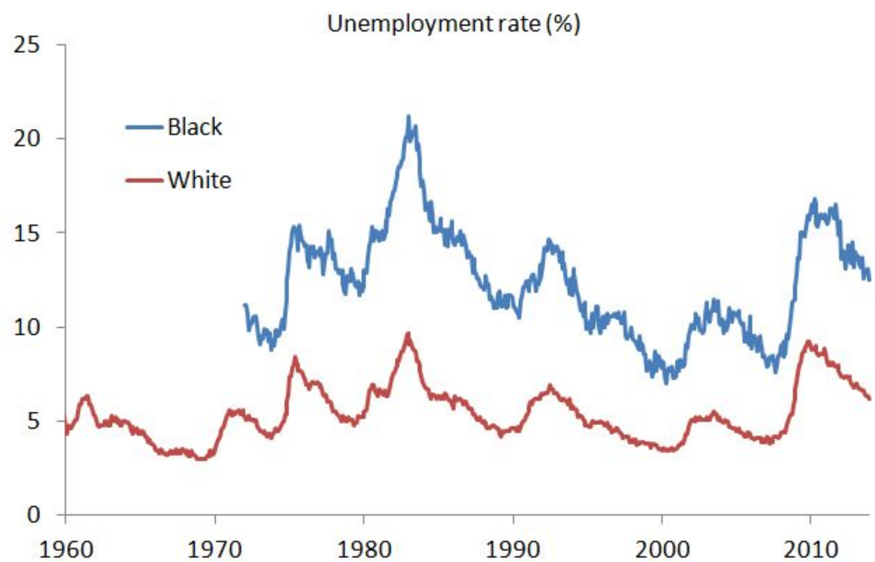
2 Stylized facts

There are two stylized facts essential to this paper. First, labor market outcomes differ by race. Figures 1 and 2 show (using BLS data) that black workers earn less and are more likely to be looking for work. Racial gaps in earnings and unemployment also exist across education levels. And across education levels, employed black workers report fewer job offers, as shown by Table 1.[5]

Second, the socioeconomic advantage of social connections is also unequally distributed by race. This is in part a mechanical consequence of racial homophily—the tendency of people to have friends of the same race. Figure 3 shows that the

[5]This table uses data from the NLSY79, described in Section 4.1. Wolpin (1992) finds that young black workers who did not go to college receive more job offers than average; these figures provide a broader view.

Figure 2: Racial inequality in unemployment



friends of black workers are more likely to be looking for work, especially at lower levels of educational attainment. The friends of black workers at all levels of educational attainment also tend to have less formal education. For example, in Wave III of the Add Health dataset[6], 21% of the friends of black respondents with exactly 12 years of education did not finish high school, compared with 14% for the friends of non-black respondents.

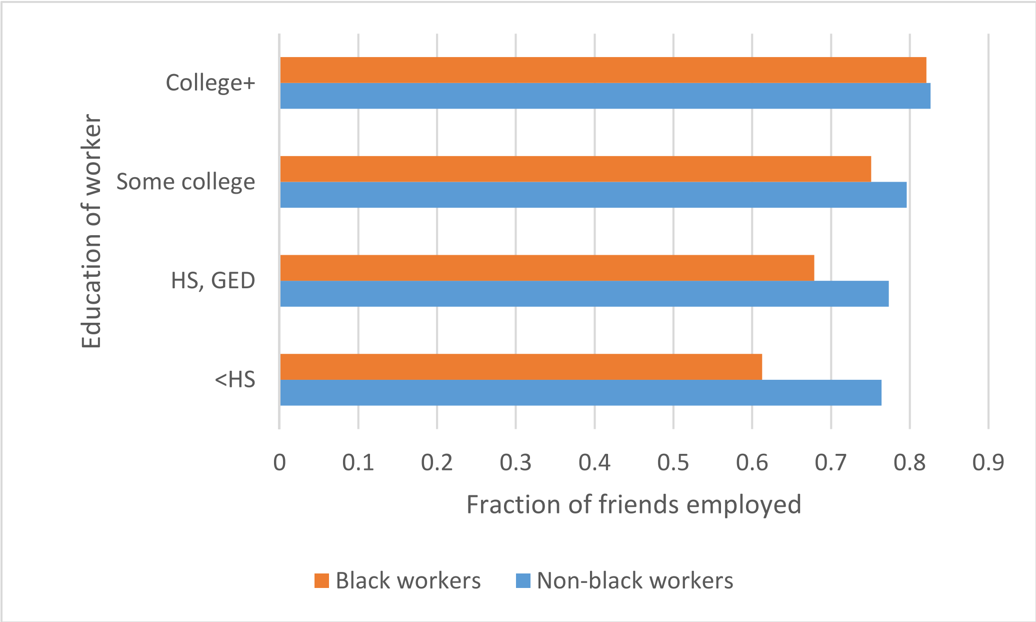
Taken together, these facts suggest that non-black workers may have an advantage in finding high-paying work through friends. To fix terminology, I will use the term "social capital" to refer to any such factors which affect jobs found through friends but not those found directly.

[6]Add Health is a longitudinal study of sample of adolescents who were nationally representative of those aged 7-12 when the study began in 1994.

Table 1: Job offers by education and race

# other offers when found job		
Years of school	Non-black	Black
<12	0.31	0.23
12	0.37	0.29
13-15	0.51	0.40
16+	0.74	0.70

Figure 3: Employment rate of friends



3 Model

Consider a group I of n workers who have completed their schooling and seek to maximize expected discounted utility from earnings. Time progresses in discrete periods $t = 1, 2, 3, \ldots$, each of which is 13 weeks long.[7] Each period, a worker may receive up to one job offer. Job offers come in two flavors: those found directly and those found through friends. The chance of receiving a direct offer is λ_{it}^{d} and the chance of receiving an offer through friends is λ_{it}^{f} (parameterized below in Equations 4 and 5). A worker who receives an offer in period t can either accept it or reject it and maintain the status quo—his current job (if employed) or unemployment (if not). This is the only choice made by workers in the model. At the end of the period, employed workers may lose their job with chance δ_{it}.

The log wage that worker i receives at time t from job j if the job was found directly is as follows:

$$w_{ij(t)}^{d} = \eta_i + \alpha_1 a_{it} + \alpha_2 a_{it}^2 + \varepsilon_{ij(t)}^{d}. \tag{1}$$

The log wage of a job found through friends is:

$$w_{ij(t)}^{f} = \beta_{\text{hcap}} \eta_i + \beta_{\text{age}} \left(\alpha_1 a_{it} + \alpha_2 a_{it}^2 \right) + \beta_0 + \beta_s s_{r(i)} + \varepsilon_{ij(t)}^{f}. \tag{2}$$

Each worker is endowed with initial marketable human capital η_i, which can be thought of as summarizing characteristics relevant to productivity such as education and ability. While η_i will be referred to as "human capital" throughout the paper, it is best thought of as i's earning power. For example, η_i will not be i's true productive capacity if discrimination means that i is paid less than her true productive capacity. In this case η_i can be thought of how much i can expect to be paid

[7] The length of each period is chosen to accomodate the fact that while in the NLSY79 data workers can report multiple jobs in a given year, nearly all report fewer than five.

under the current discriminatory regime. This is discussed further in Section 5.3 and Appendix B.

Initial human capital is distributed normally with parameters that may differ by race ($r_i \in \{0,1\}$ denotes whether i reports being black), reflecting differences in early childhood circumstances,[8] educational attainment, etc.:

$$\eta_i \sim N\left(\mu_{r(i)}, \sigma^2_{r(i)}\right). \tag{3}$$

Worker i's wage at job j grows deterministically with age, a_{it}. The expression $\left(\alpha_1 a_{it} + \alpha_2 a^2_{it}\right)$ can be thought of as additional human capital i has acquired through experience. The shock ε^d_{ij} represents how productive worker i is at firm j, and is drawn from a normal distribution with mean normalized to zero and standard deviation σ^d.

For jobs found through friends, the analogous match quality between worker and firm is $\beta_0 + \beta_s s_{r(i)} + \varepsilon^f_{ij(t)}$, where $\varepsilon^f_{ij(t)}$ is normally distributed with mean zero and standard deviation σ^f. This allows for the possibility that jobs found through friends provide better or worse productivity matches (β_0), and that this difference depends on race ($\beta_s s_{r(i)}$).

Each worker is endowed with social capital $s_{r(i)}$, which represents the average difference in social capital between black and non-black workers. Having no natural units, it is equal to one for black workers and zero otherwise. This is of course a simplification—the implications of heterogeneity in social capital within racial groups are discussed in Section 5.3 and Appendix B. Section 2 highlighted several group-level differences in social connections that we might expect to influence job

[8]For example, Aizer and Currie (2017) find that black children face higher exposure to lead, effecting higher rates of school suspension and juvenile detention.

search; the goal of this paper is to determine the degree to which such differences translate into aggregate differences in labor market outcomes. In the model, social capital affects both the arrival of new offers found through friends and the wages they promise, but not offers found directly.

A worker's wage can grow over the life cycle by remaining in the same job and gaining experience, or by moving to a job with a better productivity match quality ε_{ij}. But a worker's match quality ε_{ij} at a given job does not change over time.

When workers are unemployed, they receive unemployment insurance equal to 40% of the most recent wage up to a maximum of \$130/week (in 1979 dollars), roughly corresponding to typical US unemployment insurance laws. If a worker has never worked, as is the case when entering the labor market for the first time, the worker's initial and acquired human capital $\eta_i + \alpha_1 a_{it} + \alpha_2 a_{it}^2$ is used in place of the most recent wage.

The chance of receiving an offer directly (λ_{it}^d) and the chance of receiving an offer through friends (λ_{it}^f) are parameterized as follows:

$$\lambda_{it}^d = \left(\Lambda_0^d + \Lambda_{\text{race}}^d r_i + \Lambda_{\text{hcap}}^d \eta_i + \Lambda_{\text{age}}^d \left(\alpha_1 a_{it} + \alpha_2 a_{it}^2\right)\right) \left(\Lambda_{\text{emp}}^d\right)^{1-u_{it}} \tag{4}$$

$$\lambda_{it}^f = \left(\Lambda_0^f \left(\Lambda_0^d + \Lambda_{\text{race}}^d r_i\right) + \Lambda_{\text{hcap}}^f \eta_i + \Lambda_{\text{age}}^f \left(\alpha_1 a_{it} + \alpha_2 a_{it}^2\right) + \Lambda_s^f s_{r(i)}\right) \left(\Lambda_{\text{emp}}^f\right)^{1-u_{it}}. \tag{5}$$

The arrival of each type of offer depends on i's race r_i, initial human capital η_i, accumulated human capital $\left(\alpha_1 a_{it} + \alpha_2 a_{it}^2\right)$, an indicator u_{it} of whether i is unemployed in period t, and a constant (note that Λ_{emp}^d scales the overall chance of a direct offer when the worker is employed and $u_{it} = 0$). The arrival rate of jobs found through friends depends on all of these and also on social capital $s_{r(i)}$.

Finally, at the end of each period an employed worker's job is destroyed with chance δ_{it}:

$$\delta_{it} = \Delta_0 + \Delta_{\text{race}} r_i + \Delta_{\text{hcap}} \eta_i + \Delta_{\text{age}} \left(\alpha_1 a_{it} + \alpha_2 a_{it}^2 \right) + \Delta_f f_{ij(t)}. \tag{6}$$

The chance of losing one's job depends on $f_{ij(t)}$, an indicator of whether i's job j at time t was found through friends. Jobs found through friends may be a better match for a worker's skills, or it may be more enjoyable to work with friends. More cynically, it may be harder to fire someone who is a friend or relative of the boss. If worker i's job is not lost in period t, keeping the current job is an option in period $t+1$.

It is theoretically possible in this model that an employed worker might want to quit to unemployment. Since the chance of finding a new job changes with age, the relative value of unemployment may have increased since the worker accepted her current job offer. However, at the estimated parameter values this is never the case, and thus this possibilty is ignored: employed workers never choose to quit to unemployment.

What offers will a worker accept? Worker i's value of employment at job j with match quality ε_{ij} at time t can be written as follows. Note that the value also depend on race r_i, social capital s_i, and initial human capital η_i, but as they are time-invariant these arguments are suppressed and replaced with a dot $(\cdot)$.

$$V\left(\varepsilon_{ij(t)}, f_{ij(t)}, a_{it}, \cdot\right) = w_{ij(t)} + \rho \Bigg\{ \delta_{it} U\left(w_{ij(t)}, a_{it+1}, \cdot\right) + (1-\delta_{it}) \Bigg[V\left(\varepsilon_{ij(t)}, f_{ij(t)}, a_{it+1}, \cdot\right)$$
$$\left(\left(1-\lambda_{it}^{d}-\lambda_{it}^{f}\right) + \lambda_{it}^{d}\Phi^{d}\left(\underline{\varepsilon}^{ed}\left(\varepsilon_{ij(t)}, f_{ij(t)}, a_{it}, \cdot\right)\right) + \lambda_{it}^{f}\Phi^{f}\left(\underline{\varepsilon}^{ef}\left(\varepsilon_{ij(t)}, f_{ij(t)}, a_{it}, \cdot\right)\right)\right)$$
$$+ \lambda_{it}^{d} \int_{\underline{\varepsilon}^{ed}\left(\varepsilon_{ij(t)}, f_{ij(t)}, a_{it}, \cdot\right)}^{\infty} V\left(\varepsilon_{ij'}, 0, a_{it+1}, \cdot\right) d\phi^{d}\left(\varepsilon_{ij'}\right)$$
$$+ \lambda_{it}^{f} \int_{\underline{\varepsilon}^{ef}\left(\varepsilon_{ij(t)}, f_{ij(t)}, a_{it}, \cdot\right)}^{\infty} V\left(\varepsilon_{ij'}, 1, a_{it+1}, \cdot\right) d\phi^{f}\left(\varepsilon_{ij'}\right) \Bigg] \Bigg\}$$

First, the worker receives the wage flow, $w_{ij(t)}$. With chance δ_{it} (given by equation 6), worker i's job is lost and starts the next period (discounted by ρ) with a value of unemployment U that depends on the wage of the job lost as well as the worker's new age a_{it+1}. With chance $1-\delta_{it}$, the job is not lost. The second line contains the chance of remaining in the same job—this can happen either by receiving no offers (chance $\left(1-\lambda_{it}^{d}-\lambda_{it}^{f}\right)$) or by receiving a job offer at a firm that is below the worker's reservation match quality for offers received in employment ($\underline{\varepsilon}^{ef}$ for jobs found through friends and $\underline{\varepsilon}^{ed}$ for jobs found directly). The reservation match quality depends the current match quality and worker's age. It also depends on on how the new and old offers were found since this can affect how likely the job is to be lost. So a worker may, for example, be willing to accept a slightly lower wage for a job found through friends if it is likely to last longer. Φ^{d} and Φ^{f} are normal cumulative densities with mean zero and standard deviations σ^{d} and σ^{f}, reflecting the distribution of match quality.

The third line contains the chance and value that a direct job offer is received and the match quality $\varepsilon_{ij'}$ of the new job j' is above the reservation match quality. The fourth line contains the chance and value that an offer is received through friends and it is of acceptable match quality. Here ϕ^d and ϕ^d are normal densities of mean zero and standard deviations σ^d and σ^f, again reflecting the distribution of match quality.

The value of unemployment depends on the most recent wage $w_{ij(t)}$ at time t as well as age (when a worker is unemployed, define for convenience $w_{ij(t)}$ as their most recent wage).

$$
\begin{aligned}
&U\left(w_{ij(t)}, a_{it}, \cdot\right) = \max\left\{w_{ij(t)} + \log(0.4), \log(\$130 \times 4)\right\} + \rho \Bigg[\\
&\left(\left(1 - \lambda_{it}^d - \lambda_{it}^f\right) + \lambda_{it}^d \Phi^d\left(\underline{\varepsilon}^{ud}\left(w_{ij(t)}, a_{it}, \cdot\right)\right) + \lambda_{it}^f \Phi^f\left(\underline{\varepsilon}^{uf}\left(w_{ij(t)}, a_{it}, \cdot\right)\right)\right) U\left(w_{ij(t)}, a_{it+1}, \cdot\right) \\
&\qquad + \lambda_{it}^d \int_{\underline{\varepsilon}^{ud}\left(\varepsilon_{ij(t)}, f_{ij(t)}, a_{it}, \cdot\right)}^{\infty} V\left(\varepsilon_{ij'}, 0, a_{it+1}, \cdot\right) d\phi^d\left(\varepsilon_{ij'}\right) \\
&\qquad\qquad + \lambda_{it}^f \int_{\underline{\varepsilon}^{uf}\left(w_{ij(t)}, f_{ij(t)}, a_{it}, \cdot\right)}^{\infty} V\left(\varepsilon_{ij'}, 1, a_{it+1}, \cdot\right) d\phi^f\left(\varepsilon_{ij'}\right) \Bigg]
\end{aligned}
$$

The value of unemployment has four parts. First is unemployment insurance, which depends on the most recent wage and is described above. The second line contains the chance and value of remaining unemployed in the following period, which occurs if either no offer arrives or if one does but it is below the reservation match quality. Note that since the flow value of unemployment insurance depends on a worker's most recent wage but is invariant to how that job was found, the reservation match qualities in unemployment $\underline{\varepsilon}^{uf}$ and $\underline{\varepsilon}^{ud}$ likewise depend on the most recent wage but not its provenance. They still depend on the provenance of the new offer, however, as again this affects the chance of the job being lost. And as in Burdett (1978), a worker may prefer to remain in unemployment than take a low offer, given that the search rate is higher in unemployment. The third line contains the chance and value of an acceptable offer through direct search, and the fourth line contains the chance and value of an acceptable offer arriving through friends.

In this model there are three distinct ways in which friends can help you on the labor market:

1. More offers
2. Higher-paying offers
3. Jobs less likely to be lost[9]

The focus of this paper is whether or not social capital differs by race in a way that affects jobs found through friends through these three channels.

4 Data and preliminary analysis

4.1 Data

The model is estimated using the NLSY79, a longitudinal study of a sample of 12,686 people in the US who were nationally representative when first surveyed in 1979, at ages 14–22. In years 1982, 1994, 1996, 1998, and 2000, respondents who reported working were asked whether or not they found their job through asking friends or relatives. Specifically, those who reported looking for work when offered their current job are asked “Which of the methods on this card led to your being

[9] Loury (2006) and Brown et al. (2015) have found that jobs found through social connections last longer.

offered your job with [Name of employer]?" I identify this as a job found through friends/relatives if they marked "Contacted friends or relatives."[10]

For these jobs (up to five for each year), respondents also reported their wages, occupation, and whether they were working when they found the job. All dollar values are deflated using the CPI-U-RS published by the Bureau of Labor Statistics. Educational attainment is also reported, and the sample is restricted to include only those observations for which respondents had reached their highest level of education. The data also include sex, race, age, and employment and marital status. This paper focuses on differences in outcomes between black workers and others.

4.2 Preliminary analysis

Before estimating the full model, the main result can be roughly encapsulated in a single fixed-effects regression. Log weekly earnings are regressed on whether the job was found through friends, the interaction of this variable with race, and a variety of other controls including individual fixed effects. Selected coefficients are reported in Table 2.

[10]Other possible search methods include checking with a state employment agency, checking with a private employment agency, contacting an employer directly, placing or answering an ad, and looking in a newspaper.

Table 2: Wage premium of jobs found through friends

Fixed effects regression of log weekly earnings		
Variable	Coefficient	Std. Err.
Job found through friends	0.196	0.015
Black*(job found through friends)	-0.078	0.035
Age	0.138	0.025
Age^2	-0.002	2.48E-4
Married	-0.012	0.025
Married*male	0.173	0.034
# obs.	24,433	
# individuals	8,510	

Note: Unreported coefficients include individual fixed effects, an indicator for urban location, year dummies, and a constant.

Jobs found through friends tend to pay more, but the premium is lower for black workers. This is the main result of the paper—inequality in social capital seems to be driving part of the racial wage gap, even after controlling for individual productivity. The network wage premium does vary by occupation (see Table 10 in Appendix A), but the racial difference in network wage premium is actually slightly larger if controls for each worker's occupation are included.

How much of the racial wage gap can social capital account for? To answer this question, I predict log weekly earnings using the estimated coefficients, but setting (black*job found through friends) equal to zero for all individuals. The predicted weekly earnings, converted to 2016 dollars for convenience) are given by Table 3.

Table 3: Weekly earnings predicted by linear regression (2016 $)

	Black	Non-black	Gap
Actual	442	530	88
Predicted	442	530	88
Counterfactual	450	530	79

This suggests that inequality in social connections can explain 10% of the predicted gap in log weekly earnings.

How does this fit in with previous findings that most of the wage gap can be accounted for by premarket factors? O'Neill (1990) and Neal and Johnson (1996) find that in a regression of log wages on AFQT score and race, variation in the former is able to account for nearly all the variation in the outcome. This is interpreted as evidence that premarket factors such as human capital formation are responsible for the racial wage gap, rather than bias in hiring against black workers of equal productivity.[11] Table 11 in the Appendix shows the results of a similar exercise carried out separately for jobs found through friends and for jobs found directly. In jobs found directly, differences in AFQT can largely account (in a statistical sense) for racial inequality in earnings, as in previous work. But for jobs found through friends, the coefficient on race is large (and statistically significant), suggesting that there is racial variation in jobs found through friends that requires another explanation.

Given this result, estimating the model can fill in some important details. First, the racial discrepancy in the network wage premium could owe to better offers or

[11]To the extent that social connections are formed before entering the labor market they are a premarket factor, and variation in AFQT score may reflect both individual productivity as well as something about one's social connections, making it an imperfect way to separate an individual's characteristics from those of her social connections.

simply more offers. Distinguishing these effects requires estimating the full model. Second, the exercise above is unable to address important issues of selection. In particular, changing the offers black workers find through friends could cause some workers with multiple offers to choose a different offer, or pull some workers out of unemployment—eventualities not addressed by the reduced-form counterfactual. Finally, the full model is also better suited to distinguishing the effects of human capital form those of social capital. Suppose that jobs found through friends yield higher wages only for workers with lots of human capital. Then the racial difference in the premium for jobs found through friends estimated in Table 2 could owe entirely to a group difference in human capital, rather than social capital. By explicitly allowing the premium for jobs found through friends to depend on human capital, the full model will resolve this issue.

5 Estimation

5.1 Identification

A key feature of the data is that for each individual, multiple jobs are observed over time, some of which are found directly and some of which are found through friends. This allows individual productivity, which affects both types of offers, to be distinguished from the individual's social capital, which only affects the latter.

The 27 parameters are estimated jointly via method of simulated moments. Given a set of parameters, I first determine the reservation match quality in each state. To do so, I solve backwards assuming log utility, exogenous retirement at age 65, and a time discount factor ρ of 95% (per year). The continuous variables (initial productivity and match quality) are each approximated by a 16-point grid. Finer grids were also tried, with no appreciable effect on results.

Next, I simulate each worker's wage path from the first quarter following the end of schooling through the year 2000, when the data end. When a worker's simulated human capital or the match quality of a simulated offer lie between grid points, linear interpolation is used to determine the reservation match quality. From these complete wage paths, I censor the simulated observations to include only those worker-period observations included in the NLSY79 data set.

The censored simulated data can then be compared to the actual data through the lens of certain judiciously chosen moments and regression coefficients, which are described below and summarized in Table 4. To avoid redundancy (which would over-weight certain statistics), not every coefficient from a given regression is targeted in estimation; these coefficients are marked in Table 4 with a dagger $(\dagger)$.

First, log wages are regressed on age a_{it}, its square, an indicator f_{it} of whether the job was found through friends, and the interaction of this indicator with race r_i and age. This regression also includes individual dummy variables $\{d_i\}$; the estimated coefficients on these indicator variables (fixed effects) are denoted $\hat{\eta}_i$. The five coefficients not including the fixed effects help identify the parameters $\{\alpha_1, \alpha_2\}$ from Equation 1, which determine the growth of human capital with age and its square. They also help identify the three parameters $\left\{\beta_0, \beta_{\text{age}}, \beta_s\right\}$ in equation 2, which determine the distribution of wages of jobs found through friends. The final parameter in this equation, β_{edu}, is pinned down by the coefficient on the estimated fixed effect $\hat{\eta}_i$ in a separate regression of log wages on jobs found through friends that also includes a constant, race r_i, and age a_{it}.

A regression of the estimated fixed effect $\hat{\eta}_i$ on a constant, race, age, and its square identifies the means of the initial human capital distribution $\left\{\mu_{r(i)}\right\}$ from Equation 3 and also helps identify the dependence of human capital on experience. The standard deviations of the fixed effect for non-black and black workers identify the dispersion of initial human capital $\left\{\sigma_{r(i)}\right\}$, and the standard deviations of wages less the fixed effect $(w_{it} - \hat{\eta}_i)$ for jobs found directly and through friends identify the within-worker dispersion of offers $\left\{\sigma_f, \sigma_d\right\}$.

A regression of unemployment u_{it} on race, the estimated fixed effect $\hat{\eta}_i$, age, and a constant targets the parameters of the job destruction rate $\left\{\Delta_{\text{race}}, \Delta_{\text{hcap}}, \Delta_{\text{age}}, \Delta_0\right\}$ in Equation 6. A similar regression of log tenure at the current job on the same regressors targets the arrival rate parameters $\left\{\Lambda^d_{\text{race}}, \Lambda^d_{\text{hcap}}, \Lambda^d_{\text{age}}, \Lambda^d_0\right\}$ and $\left\{\Lambda^f_{\text{race}}, \Lambda^f_{\text{hcap}}, \Lambda^f_{\text{age}}, \Lambda^f_0, \Lambda^f_s\right\}$ in Equations 4 and 5 as well as the dependence of job loss on provenance Δ_f, and a regression of f_{it} (an indicator of whether the current job was found through friends) on these regressors distinguishes the arrival rate parameters for direct offers from those for offers through friends.

An indicator that i was employed when the current job was found is averaged for jobs found directly and those through friends to distinguish on-the-job search from search in unemployment for both sources of offers, parameterized by Λ^d_{emp} and Λ^f_{emp}.

5.2 Procedure

Practically, estimation proceeds as follows. First, a vector of targeted statistics in Table 4 is calculated using the actual data. Then, starting values of the parameters

Table 4: Targeted statistics

	Dependent variable	Included independent variables								
		1	r_i	a_{it}	${a_{it}}^2$	f_{it}	$f_{it} \cdot r_i$	$f_{it} \cdot a_{it}$	$\{d_i\}$	$\hat{\eta}_i$
(1)	Log wage w_{it}			$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark^{\dagger}$	
(2)	w_{it}, job through friends	$\checkmark^{\dagger}$	$\checkmark^{\dagger}$	$\checkmark^{\dagger}$						$\checkmark$
(3)	Estimated fixed effect $\hat{\eta}_i$	$\checkmark$	$\checkmark$	$\checkmark^{\dagger}$	$\checkmark^{\dagger}$					
(4)	Std. dev. of $\hat{\eta}_i$	$\checkmark$	$\checkmark$							
(5)	Std. dev. of $(w_{it} - \hat{\eta}_i)$	$\checkmark$				$\checkmark$				
(6)	Unemployment u_{it}	$\checkmark$	$\checkmark$	$\checkmark$						$\checkmark$
(7)	Log tenure at current job	$\checkmark$	$\checkmark$	$\checkmark$		$\checkmark$				$\checkmark$
(8)	Job found through friends f_{it}	$\checkmark$	$\checkmark$	$\checkmark$						$\checkmark$
(9)	Employed when found job	$\checkmark$				$\checkmark$				

†Coefficients not targeted in estimation, to avoid redundancy.

are chosen from preliminary regressions. The model is simulated, and a vector of targeted statistics is calculated from the simulated outcomes. The difference between the the vector of true statistics and the vector of simulated statistics is the sum of squared elementwise differences, each divided by the magnitude of the true statistic.

$$\rho\left(v^{\text{true}}, v^{\text{sim}}\right) = \sum_i \frac{\left(v_i^{\text{true}} - v_i^{\text{sim}}\right)^2}{|v_i^{\text{true}}|} \tag{7}$$

An algorithm[12] searches for the set of parameters which, when simulated, yield a vector of statistics v^{sim} that minimizes this distance $\rho\left(v^{\text{true}}, v^{\text{sim}}\right)$.

Finally, confidence intervals are constructed by bootstrapping. The set of workers is resampled with replacement to generate a synthetic sample, and the parameters are reestimated using the synthetic sample. The baseline and counterfactual

[12]Specifically, MATLAB's fminsearch is used, which employs the Nelder-Mead simplex algorithm. This was supplemented with the Artelys Knitro 10.3 solver.

simulations are run using the new parameters and synthetic sample. This is repeated 500 times to generate confidence intervals.

5.3 Caveats

This section concludes with some important caveats. There are reasons to suspect this paper's estimate of the role of social capital may be an underestimate. First, if there is racal bias in direct hiring, η_i will not represent i's productivity.[13] This alone does not affect the main results, but if there is more racial bias or discrimination against black workers in direct hiring than in hiring through friends, then the wages of black workers will understate their human capital. In this case, the racial difference in social capital will be larger than estimated.

Second, if there is within-race heterogeneity in social capital that is correlated with human capital, then wages of jobs found through friends may appear to depend on human capital more than they do. This would lead to an underestimate of the effect of increasing social capital.[14] For these reasons, this paper's results should be seen as conservative. Appendix B examines both cases in further detail.

[13] Bertrand and Mullainathan (2004) find that employers are less likely to respond to applications with names indicating a black applicant, though Neal and Johnson (1996) argue that most of the wage gap between black workers and others is accounted for by differences in skill as measured by the AFQT.

[14] If social capital were negatively correlated with human capital, this paper would overestimate the importance of social capital. However, all available evidence contradicts this possibility: self-reported sociability is positively correlated with AFQT score and wages, and the network wage premium grows with education.

6 Results

6.1 Parameter estimates

Table 5 presents the estimates of the parameters related to the distribution of human capital and its growth with work experience. The estimated group difference in mean initial human capital reflects inequality in educational attainment as well as other factors mentioned in Section 3.

Table 5: Human capital parameter estimates

		Estimate	95% CI
Mean initial human capital, non-black	$\mu_{\text{non-black}}$	1.032	[1.016, 1.042]
Mean initial human capital, black	μ_{black}	0.854	[0.834, 0.862]
Std. dev. initial human capital, non-black	$\sigma_{\text{non-black}}$	0.181	[0.180, 0.186]
Std. dev. initial human capital, black	σ_{black}	0.086	[0.084, 0.091]
Dependence of direct wages on age	α_1	0.1253	[0.1233, 0.1267]
Dependence of direct wages on age squared	α_2	-1.25 E-3	[-0.0013, -0.0012]

The other parameters that determine the wage distribution of offers are given in Table 6. Recall that the dependence of offers found directly on human capital is normalized to one, so there are no analogues of β_{hcap}, β_{exp}, or β_0 for jobs found directly. The wages of jobs found through friends are higher but have slightly lower variance, suggesting that they are better matched to worker's skills (though Loury (2006) argues that lower-paying jobs that are found through friends as a last resort can be important as well). Importantly, the racial difference in social capital does seem to affect wage offers, even after controlling for the fact that the wages of jobs found through friends may depend on human capital differently than those of jobs found directly.

Table 6: Wage offer distribution parameter estimates

		Estimate	95% CI
Wage, jobs found through friends	β_0	1.074	[1.065, 1.100]
Dependence on initial human capital η_i, jobs found through friends	β_{hcap}	0.602	[0.597, 0.617]
Dep. on age, jobs found through friends	β_{age}	0.973	[0.969, 0.984]
Dep. on racial difference in social capital $s_{r(i)}$, jobs found through friends	β_s	-0.267	[-0.270, -0.257]
Standard deviation of shocks, direct offers	σ_d	0.354	[0.350, 0.361]
Standard deviation of shocks, offers found through friends	σ_f	0.336	[0.332, 0.342]

Table 7 presents the offer receipt parameter estimates. At a given level of human capital, black workers receive more offers through friends than non-black workers. This is perhaps unexpected given Table 1, but consistent with Wolpin (1992).

Table 7: Offer receipt parameter estimates

		Estimate	95% CI
Base chance of receiving an offer directly in unemployment	Λ_0^d	2.778	[2.611, 2.816]
Base chance (as a fraction of Λ_0^d), jobs found through friends	Λ_0^f	0.276	[0.272, 0.281]
Difference for black workers, jobs found directly	Λ_{race}^d	-4.15 E-3	[-4.22 E-3, -4.12 E-3]
Difference for black workers, jobs found through friends	Λ_{race}^f	-4.14 E-4	[-4.24 E-4, -4.14 E-4]
Dependence on initial human capital η_i, jobs found directly	Λ_{hcap}^d	3.20 E-4	[3.18 E-4, 3.24 E-4]
Dependence on initial human capital η_i, jobs found through friends	Λ_{hcap}^f	5.15 E-3	[5.10 E-3, 5.23 E-3]
Dependence on age, jobs found directly	Λ_{age}^d	0.314	[0.309, 0.319]
Dependence on age, jobs found through friends	Λ_{age}^f	1.01 E-2	[1.01 E-2, 1.03 E-2]
Rate when employed, jobs found directly	Λ_{emp}^d	3.25 E-2	[3.02 E-2, 3.30 E-2]
Rate when employed, jobs found through friends	Λ_{emp}^f	7.86 E-2	[7.77 E-2, 8.07 E-2]
Dependence of arrival rate on racial difference in social capital	Λ_s	0.102	[0.102, 0.105]

Table 8 gives the estimates of the job destruction parameters. Human capital lowers the rate of job destruction, and black workers are more likely to lose their jobs.

Table 8: Job destruction parameter estimates

		Estimate	95% CI
Base destruction rate	Δ_0	0.0238	[0.0236, 0.0242]
Difference for black workers	Δ_{race}	0.001	[9.49 E-4, 9.77 E-4]
Dependence on initial human capital η_i	Δ_{hcap}	-0.005	[-5.32 E-3, -5.06 E-3]
Dependence on age	Δ_{age}	-0.006	[-6.53 E-3, -6.30 E-3]
Dependence on how job was found f_{it}	Δ_{f}	-3.06 E-4	[3.12 E-4, 3.04 E-4]

6.2 Simulated statistics

With the parameters estimated, it is possible to simulate the model in both baseline and counterfactual scenarios. Table 9 presents the values of some key moments of interest for both simulated and actual data, as well as a counterfactual in which there is no racial difference in social capital (i.e., s_i is set to zero for all workers). Estimates of the full list of 27 targeted moments and regression coefficients from Table 4 are in Table 12 in Appendix A. The model does quite well at matching earnings and the frequency with which accepted job offers are found through friends, while the simulated unemployment rates are lower than they are in the data, especially for non-black workers.

Table 9: Key moments

		Data	Simulated	Counterfactual	Difference
% jobs found through friends	Non-black	23.7	20.3	-	-
			[18.8, 25.4]	-	-
	Black	26.2	21.4	37.5	16.1
			[17.9, 26.6]	[32.3, 39.9]	[11.8, 16.3]
Weekly earnings of employed (2016 $)	Non-black	511	501	-	-
			[455, 534]	-	-
	Black	442	435	443	8.35
			[400, 477]	[411, 480]	[-0.1, 18.2]
Unemployment rate, %	Non-black	6.15	3.68	-	-
			[2.47, 5.03]	-	-
	Black	8.36	7.92	5.66	-2.25
			[6.08, 11.10]	[3.96, 8.20]	[-3.47, -1.77]

Note: The counterfactual sets the social capital of black workers equal to that of non-black workers; since this doesn't affect non-black workers, their outcomes are unchanged and thus marked with a dash. Bootstrapped 95% confidence intervals are given in brackets. The rightmost column gives the difference between the counterfactual simulation and the baseline simulation.

The counterfactual exercise is similar to that undertaken in Table 3. Specifically, the racial difference in social capital is set to zero to determine how much of the wage and unemployment gaps it can account for. The main finding is larger than that of Table 3 but qualitatively similar: racial differences in jobs found through friends can account for 13% of the simulated racial gap in earnings of employed workers (95% confidence interval: [0.0%, 32.9%]). The simulated racial unemployment gap narrows by two percentage points, though this result should be interpreted with caution given that the unemployment rates were not as well matched by the simulation in the baseline scenario. Nonetheless, this suggests that social connections affect labor market inequality on both extensive and intensive margins. Since the job finding rate for black workers actually declines in the counterfactual, these improvements in labor market outcomes owe entirely to the offer distribution from which jobs found through friends are drawn.

7 Conclusion

This paper is the first to estimate the contribution of social capital inequality to the aggregate wage gap between black workers and others in the US. I find that at least a tenth of the gap can be explained by differences in jobs found through friends—a conservative estimate which may be biased downward by within-group correlation between social capital and human capital (discussed in Section 5.3). More broadly, these results demonstrate the importance of considering social capital in the study of wage inequality between groups.

The policy implications of these results may not be immediately obvious. Even if friends are important for finding work, can policy force people to make different friends? Perhaps not directly, but if social connections made early in life persist then neighborhood and school segregation may be important in determining labor market outcomes. Furthermore, this paper does have something to say about interventions that do not directly alter social networks. The main result of this paper is that a significant fraction of group inequality in labor market outcomes owes to heterogeneity other than worker productivity that prevents certain workers from finding jobs that reward their skills. Accordingly, a policy that affects the arrival rate of offers such as affirmative action in hiring may be able to counteract the racial imbalance in job-finding.

References

Aizer, A. and Currie, J. (2017). Lead and Juvenile Delinquency: New Evidence from Linked Birth, School and Juvenile Detention Records. Working Paper 23392, National Bureau of Economic Research. DOI: 10.3386/w23392.

Altonji, J. G. and Blank, R. M. (1999). Race and Gender in the Labor Market. *Handbook of Labor Economics*, 3:3143–3259.

Arrow, K. J. and Borzekowski, R. (2004). Limited Network Connections and the Distribution of Wages. FEDS Working Paper 2004-41, Board of Governors of the Federal Reserve System.

Becker, G. S. and Tomes, N. (1979). An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. *Journal of Political Economy*, 87(6):1153–1189.

Bertrand, M. and Mullainathan, S. (2004). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, 94(4):991–1013.

Brown, M., Setren, E., and Topa, G. (2015). Do Informal Referrals Lead to Better Matches? Evidence from a Firm's Employee Referral System. *Journal of Labor Economics*, 34(1):161–209.

Burdett, K. (1978). A Theory of Employee Job Search and Quit Rates. *The American Economic Review*, 68(1):212–220.

Calvó-Armengol, A. and Jackson, M. O. (2004). The Effects of Social Networks on Employment and Inequality. *The American Economic Review*, 94(3):426–454.

Coate, S. and Loury, G. C. (1993). Will Affirmative-Action Policies Eliminate Negative Stereotypes? *The American Economic Review*, 83(5):1220–1240.

Durlauf, S. N. (1996). A theory of persistent income inequality. *Journal of Economic Growth*, 1(1):75–93.

Fang, H. and Moro, A. (2011). Theories of Statistical Discrimination and Affirmative Action: A Survey. In Benhabib, J., Jackson, M. O., and Bisin, A., editors, *Handbook of Social Economics*, volume 1A, pages 133–200. Elsevier.

Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, 78(6):1360–1380.

Green, G. P., Tigges, L. M., and Diaz, D. (1999). Racial and Ethnic Differences in Job-Search Strategies in Atlanta, Boston, and Los Angeles. *Social Science Quarterly*, 80(2):263–278.

Holzer, H. J. (1987). Informal Job Search and Black Youth Unemployment. *The American Economic Review*, 77(3):446–452.

Ioannides, Y. M. and Datcher Loury, L. (2004). Job Information Networks, Neighborhood Effects, and Inequality. *Journal of Economic Literature*, 42(4):1056–1093.

Loury, G. C. (1981). Intergenerational Transfers and the Distribution of Earnings. *Econometrica*, 49(4):843–867.

Loury, L. D. (2006). Some Contacts Are More Equal than Others: Informal Networks, Job Tenure, and Wages. *Journal of Labor Economics*, 24(2):299–318.

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

Montgomery, J. D. (1991). Social Networks and Labor-Market Outcomes: Toward an Economic Analysis. *The American Economic Review*, 81(5):1408–1418.

Montgomery, J. D. (1992). Social Networks and Persistent Inequality in the Labor Market. *Working Paper*.

Neal, D. A. and Johnson, W. R. (1996). The Role of Premarket Factors in Black-White Wage Differences. *Journal of Political Economy*, 104(5):869–895.

O'Neill, J. (1990). The Role of Human Capital in Earnings Differences Between Black and White Men. *The Journal of Economic Perspectives*, 4(4):25–45.

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

Schmutte, I. M. (2016). Labor markets with endogenous job referral networks: Theory and empirical evidence. *Labour Economics*, 42:30–42.

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

Wolpin, K. I. (1992). The Determinants of Black-White Differences in Early Employment Careers: Search, Layoffs, Quits, and Endogenous Wage Growth. *Journal of Political Economy*, 100(3):535–560.

A Ancillary tables

Table 10 shows that the conditional wage premium for jobs found through friends differs across occupations, as does the frequency with which jobs are found through friends. The wage premia are calculated by a fixed effects regression akin to that in Table 2, but also including dummies for the interaction of occupation with whether a job was found through friends.

Table 10: Job market importance of friends by occupation

Occupation	# obs.	% found through friends	Wage premium (log pts)	Std. error
Professional and Technical	2,086	17.5	0.350	0.040
Sales	3,396	19.7	0.209	0.033
Operatives ex. Transport	892	25.9	0.201	0.056
Clerical and Kindred	2,055	23.8	0.201	0.038
Craftsmen and Kindred	1,768	26.2	0.181	0.043
Transport Equipment Operatives	1,299	24.4	0.134	0.049
Managers and Admin. ex. Farm	833	18.2	0.083	0.065
Farm Laborers and Foremen	3,073	20.6	0.064	0.038

Table 11 shows the results of log wages of jobs found directly and jobs found through friends separately regressed on age-adjusted, normalized AFQT score as well as controls for age, age squared, and year. The results for jobs found directly are similar to those of O'Neill (1990) and Neal and Johnson (1996), but the coefficient on race is much larger for jobs found through friends.

Table 11: Race, AFQT, and jobs found through friends

	Regression of log weekly earnings			
	Found directly		Found through friends	
Variable	Coefficient	Std. Err.	Coefficient	Std. Err.
AFQT	0.1601	0.0076	0.1479	0.0117
$AFQT^2$	0.0258	0.0068	0.0018	0.0106
Black	-0.0266	0.0182	-0.1155	0.0282
Age	0.2379	0.0179	0.3131	0.0220
Age^2	-0.0032	2.62 E-4	-0.0045	3.38 E-4
# obs.	19,052		5,381	

Note: Unreported coefficients include year dummies and a constant.

Table 12: Targeted statistics, data vs. simulated

Dependent variable	Indep. var./subset	Data	Simulated	95% CI
Log wage w_{it}	p_{it}	0.0223	0.0799	[0.0729, 0.0945]
	${p_{it}}^2$	-0.0014	0.0049	[0.0034, 0.0077]
	f_{it}	4.0607	3.2842	[3.1847, 3.3991]
	$f_{it} \cdot r_i$	-0.2284	-0.2819	[-0.2993, -0.1570]
	$f_{it} \cdot p_{it}$	0.0037	0.0042	[-0.0032, 0.0061]
w_{it}, job through friends	$\hat{\eta}_i$	0.9391	0.2247	[0.1094, 0.2247]
Estimated fixed effect $\hat{\eta}_i$	1	3.8972	3.6090	[3.3835, 3.7194]
	r_i	-0.1746	-0.1094	[-0.1568, -0.0808]
Std. dev. of $\hat{\eta}_i$	Non-black	0.6629	0.6005	[0.5379, 0.6796]
	Black	-0.0968	-0.0714	[-0.1013, 0.0498]
Std. dev. of $(w_{it} - \hat{\eta}_i)$	Direct	0.5134	0.7078	[0.6572, 0.7716]
	Through friends	0.4777	0.6479	[0.6325, 0.7694]
Unemployment u_{it}	1	0.1439	0.0433	[0.0308, 0.0528]
	r_i	0.0283	0.0337	[0.0220, 0.0684]
	$\hat{\eta}_i$	-0.0273	-0.0045	[-0.0121, -0.0005]
	p_{it}	-0.0026	-0.0036	[-0.0044, -0.0026]
Log tenure at current job	1	2.0226	0.9249	[0.5802, 1.1527]
	r_i	-0.1204	-0.0271	[-0.1293, 0.0666]
	$\hat{\eta}_i$	0.6162	0.1901	[0.1347, 0.2838]
	p_{it}	0.0397	-0.1281	[-0.1380, -0.1261]
	f_{it}	0.0856	0.1009	[0.0422, 0.1595]
Job found through friends f_{it}	1	0.3682	0.2223	[0.2582, 0.3967]
	r_i	0.0251	0.0139	[-0.0218, 0.0258]
	$\hat{\eta}_i$	-0.0348	-0.0056	[-0.0481, -0.0160]
	p_{it}	-0.0049	-0.0026	[-0.0040, -0.0011]
Employed when found job	1	0.4439	0.7472	[0.7307, 0.7624]
	f_{it}	0.4306	0.7200	[0.6994, 0.7281]

B Potential sources of bias

This section examines potential sources of bias in the estimation of key coefficients. Under the most likely scenarios, this paper underestimates the racial difference in social capital and its effect on wages.

B.1 Within-race correlation between social and human capital

Suppose that the log wage of a job found through friends is:

$$w^f_{ij(t)} = \beta_0 + \beta_{\text{hcap}}\eta_i + \beta_{\text{age}}\left(\alpha_1 a_{it} + \alpha_2 a^2_{it}\right) + \beta_s s_i + \varepsilon^f_{ij(t)}. \tag{8}$$

This is essentially the same wage equation as Eq. 2, but social capital can now differ by individuals as well as race. Without loss of generality, decompose i's social capital into an average for i's race as well as an idiosyncratic component ξ_i:

$$s_i = r_i + \xi_i.$$

Using the orthogonality of $\bar{s}_{r(i)}$ with $\varepsilon_{ij(t)}$, ξ_i, and the age terms (and using $\langle\cdot,\cdot\rangle$ to denote covariance), this yields

$$\beta_s = \frac{\left\langle r_i, w^f_{ij(t)}\right\rangle - \beta_{\text{hcap}}\langle\eta_i, r_i\rangle}{\langle r_i, r_i\rangle}.$$

Similarly,

$$\beta_{\text{hcap}} = \frac{\left\langle \eta_i, w^f_{ij(t)}\right\rangle - \beta_s\langle\eta_i, r_i\rangle - \beta_s\langle\eta_i, \xi_i\rangle}{\langle\eta_i, \eta_i\rangle}.$$

Combining these yields

$$\beta_s = \frac{\left\langle r_i w^f_{ij(t)}\right\rangle\langle\eta_i, \eta_i\rangle - \langle\eta_i, r_i\rangle\left\langle \eta_i, w^f_{ij(t)}\right\rangle}{\langle r_i, r_i\rangle\langle\eta_i, \eta_i\rangle - \langle\eta_i, r_i\rangle^2 - \langle\eta_i, r_i\rangle\langle\eta_i, \xi_i\rangle}. \tag{9}$$

Compare this to the estimate if we simply project $w_{ij(t)}$ on a constant, r_i, η_i, a_{it}, and a_{it}^2 :

$$\hat{\beta}_s = \frac{\left\langle r_i w^f_{ij(t)} \right\rangle \langle \eta_i, \eta_i \rangle - \langle \eta_i, r_i \rangle \left\langle \eta_i, w^f_{ij(t)} \right\rangle}{\langle r_i, r_i \rangle \langle \eta_i, \eta_i \rangle - \langle \eta_i, r_i \rangle^2}. \tag{10}$$

Assume $\langle \eta_i, r_i \rangle$ and $\beta_s > 0$ (this is flipping the race dummy from its interpretation in the rest of the paper, but makes the comparison much easier). If human capital η_i is correlated with within-race social capital ξ_i as well as race r_i, then the true value of β_s will be larger than an estimate which ignores ξ_i (and this is assuming we knowη_i). The intuition is that correlation between ξ_i and η_i inflates the apparent dependence of wages on human capital, resulting in an underestimate of the role of social capital.

B.2 Discrimination in direct hiring

Now suppose there is discrimination in direct hiring as well as in jobs found through friends, such that

$$w^d_{ij(t)} = \left(\eta_i - \kappa^d r_i\right) + \alpha_1 a_{it} + \alpha_2 a_{it}^2 + \varepsilon^d_{ij(t)}. \tag{11}$$

$$w^f_{ij(t)} = \beta_{\text{hcap}} \left(\eta_i - \kappa^f r_i\right) + \beta_{\text{age}} \left(\alpha_1 a_{it} + \alpha_2 a_{it}^2\right) + \beta_0 + \beta_s s_i + \varepsilon^f_{ij(t)}. \tag{12}$$

Here black workers wages of jobs found through friends are reduced by κ^d and those of jobs found through friends are reduced by κ^f. If we estimate individual fixed effects from the direct wage equation, we will be getting an estimate of $\eta_i - \kappa^d r_i$ rather than true human capital η_i. Now, notice that Equation 12 can be rewritten

$$
\begin{aligned}
w_{ij(t)}^{f} = {} & \beta_{\text{hcap}}\left(\eta_i - \kappa^d r_i\right) + \beta_{\text{age}}\left(\alpha_1 a_{it} + \alpha_2 a_{it}^2\right) \\
& + \beta_0 + \beta_s s_i + \beta_{\text{hcap}}\left(\kappa^d - \kappa^f\right) r_i + \varepsilon_{ij(t)}^{f}.
\end{aligned}
$$

So if we project $w_{ij(t)}^{f}$ on $\eta_i - \kappa^d r_i$, the age terms, a constant, and r_i, instead of β_s we will get as the estimated coefficient on race

$$
\beta_s + \beta_{\text{hcap}}\left(\kappa^d - \kappa^f\right). \tag{13}
$$

So if $\beta_s < 0$, $\beta_{\text{hcap}} > 0$, and there is more discrimination in direct hiring than hiring through friends ($\kappa^d > \kappa^f$), we will underestimate the magnitude of β_s.