Who Works Where and Why? Parental Networks and the Labor Market

Shmuel San*

November 4, 2020

Click here for the latest version

Abstract

Social connections are valuable for workers entering the labor market, both because they might increase the likelihood of knowing about a job opening at a specific firm and because they may increase the likelihood of being hired, conditional on knowing about an opening. Using data from Israel and relying on identifying variation from the timing of job movements of parents' coworkers, I find that workers are three to four times more likely to find employment in firms where their parents have professional connections than in otherwise similar firms. I use the same variation to structurally estimate a matching model of the labor market with search frictions, and find that connections double the probability of meeting. The estimated value of one additional meeting with a connected firm is 3.7% of the average wage, with around 1/3 of the increase coming from the direct value of a connection. Connections matter for inequality; I find that the wage gap between Arabs and Jews decreases by 12% when equalizing the groups' connections, but increases by 56% when prohibiting the hiring of connected workers. These seemingly opposing results are explained by the fact that Arabs have connections to lower-paying firms but use their connections more extensively.

*Department of Economics, New York University, 19 W. 4th Street, 6FL New York, NY 10012. Email: muly.san@nyu.edu. Special thanks to my advisors, Christopher Flinn, Alfred Galichon, Petra Moser, and Martin Rotemberg, for guidance and support. For their helpful comments, I thank Milena Almagro, Zlil Aloni, Jaime Arellano-Bover, Pierre-André Chiappori, Tim Christensen, Christopher Conlon, Michael Gilraine, Walker Hanlon, Elena Manresa, Konrad Menzel, Maor Milgrom, Bernard Salanié, Sharon Traiberman, Miguel Uribe, and Daniel Waldinger as well as seminar participants at NYU.

1 Introduction

That some firms pay workers with similar skills differently is well documented (Abowd et al. 1999; Mortensen 2003; Card et al. 2018). Much less is known about why some workers find higher-paying jobs than comparable peers. As many, if not most, jobs are obtained through social contacts (Topa 2011), a natural answer to this question concerns differences in social networks. This article studies the role of social connections in explaining where people find their first job. I focus on one particular mechanism: firms where workers have connections through their parents.

This question has important implications for between-group inequality. Differences in the quality of labor-market ties can partly explain pay gaps between groups. This paper focuses on Israel, where there is a big pay gap between the two major ethnic groups, Jews and Arabs. Using a matched employer-employee dataset linked to the Israeli population registry, I show that Arabs have parental connections to lower-paying firms and I study the importance of that mechanism in explaining the pay gap.

I distinguish strong and weak parental connections. Strong (direct) connections are connections between employees and firms where their parents have worked. Weak (indirect) connections are connections between employees and firms where their parents' past coworkers have worked. The first part of the paper studies the reduced-form relationship between parental connections and first-job assignments as well as wages. I start with the impact of connections on a worker's first job.

To identify the effect of weak connections, I leverage the timing of both the formation and destruction of links. In particular, I compare the likelihood of working in a firm where the employee had active links in the labor-market entry year ("weak connections") with the likelihood of working in a firm where the contact had left a short time before or afterwards ("phantom connections"). I show that firms with weak and phantom connections are similar on a variety of characteristics such as sector and location.

I find that workers are 3.7 times more likely to find employment in firms with (real) weak parental connection than in phantom connected firms. Workers' probability of starting at a particular firm discretely falls the year after the link is destroyed.

To check for the possibility that estimated effects reflect endogenous separations, I estimate the effects using two exogenous causes of separation, coworkers' deaths and retirements. Specifically, I compare the probability of working at firms in which parents' coworkers died or retired after the labor-market entry year and firms in which contacts died or retired a few years before.[1] These estimates are similar in magnitude to the benchmark result, with odds

[1] See Azoulay et al. (2010) and Jager (2016) for early use of death for exogenous variation in networks.

ratios of 2.6 and 3.9 for the "death" and "retirement" connections, respectively. Likewise, to check the potential difference in employment trends between weak and phantom connections, I perform a placebo test, assigning a worker's connections to a random worker with similar observable characteristics. I find no hiring differences between phantom and real connections of a placebo worker.

I also study the heterogeneity of the effect as a function of the connections' and workers' characteristics. Connections are more effective if formed at smaller firms, for more extended periods, and more recently. Notably, connections are also stronger if the child, parent, and parent's coworker share characteristics such as gender or ethnicity. Likewise, the effect is stronger for males, Arabs, and less-educated workers.

I end the first part of the paper by studying the relationship between social connection and salary. Weak connections are associated with 1.4 percent higher wages than phantom connections. However, this analysis does not identify the causal effect of social connections on wages since it ignores selection: without connections, a hired connected worker may have counterfactually not received an offer at all instead of a different salary.[2] The model addresses this issue by jointly studying questions of matching and wage-setting.

In particular, I focus on two mechanisms through which social connections can be valuable. Firstly, they might alleviate search frictions by improving the information flow about a job opening at a specific firm and potential job seekers. Secondly, conditional on that mutual knowledge, they may increase the probability of a match between the job seeker and the firm. To study the importance of the two mechanisms, in the second part of the paper, I build and estimate a two-sided matching model of the labor market with search frictions. Matching takes place in two stages. In the first stage, workers and firms meet randomly, and the probability of meeting can vary as a function of connections. In the second stage, workers and firms that have met choose their optimal (stable) match, based on the utility they obtain from the match, which also might be affected by social connections. To separately identify the two mechanisms, I use two distinct types of information: where individuals end up working, and how much they are paid.

I estimate the model using a simulation-based method that allows for rich and flexible value functions. Getting the model's equilibrium matching and wages is computationally feasible due to the sparsity of the data resulting from the model's first stage, which restricts the set of potential matches. I estimate the parameters of the model using a novel update mapping that "inverts" the information on the observed matches and wages into the meeting and value parameters, in the spirit of the BLP contraction mapping (Berry et al. 1995).[3]

[2] Unlike the matching question where the outcome (working or not) is observed for each worker-firm combination, the outcome of the wage-setting question is only observed if the worker is hired by the firm.

[3] To the best of my knowledge, this is the first use of a "BLP-style" update mapping that estimate two

The model estimates suggest that both the "search frictions" and "match value" mechanisms are important in explaining why parental connections increase the probability of working in a firm. Weak connections increase the meeting probability by a factor of 2 and the match utility by 2.8 percent.

Using the model, I evaluate two sets of counterfactuals. Both counterfactuals rely on the identifying assumption that the causal impact of connections is the excess effect of real connections relative to phantom connections ("causal connections"). The first set of counterfactuals I evaluate is the willingness to pay for meetings and connections. I find that the average value of one additional meeting with an unconnected firm is 2.2% of new workers' average wages. On the other hand, isolating only the match quality mechanism by adding a causal weak connection to a random existing meeting increases the wage by 1.5% the average wage. Combining the two mechanisms, the value of a new meeting with causal weak connections is 3.7% of the average wage. 84% of the wage effect can be attributed to workers moving to the new connected firm, whereas the remaining 16% is due to improving workers' choice set, without changing their job.

In the second set of counterfactual exercises, I check how much of the pay gap between Jews and Arabs in Israel is due to Jews having parental connections to higher-paying firms. I find that, if Arabs and Jews had the same quantity and quality of connections, the ethnic wage gap would decrease by 12% compared with the actual gap. However, when prohibiting the hiring of connected workers, the ethnic pay gap would increase by 56%. Two opposing forces are at play in these two scenarios. On the one hand, Arabs have connections to lower-paying firms than Jews. Therefore, equalizing connections provides Arabs with better connections, which reduces the pay gap. On the other hand, Arabs rely more heavily on connections. Prohibiting the use of connections increases the gap as it hurts Arabs more than Jews.

In this paper, I focus on one particular channel for indirect connections: the network of parents' past coworkers. Existing literature studying parental connections finds that direct links (parent worked at the firm) increase the child's probability of working there; however, there is less evidence for the impact of indirect parental connections (Corak and Piraino 2011; Kramarz and Skans 2014; Plug et al. 2018). The positive effect I find for the channel of parent's past coworkers network compared to other channels of indirect networks (e.g. parents of high-school classmates or high-school classmates of one's parents) is consistent with a literature showing the importance of coworker networks for labor market outcomes (Granovetter 1973; Cingano and Rosolia 2012; Caldwell and Harmon 2018).[4]

sets of unobserved parameters using two sets of data points.

[4] I also find that the effect of connections decays over time, which explains why links formed a long time

Following the seminal study by Abowd et al. (1999), a large body of literature has documented the importance of firms in determining the wage distribution (Card et al. 2013, 2018; Song et al. 2019; Bonhomme et al. 2019). Firms play a significant role in explaining the pay gaps between different groups of workers, including males and females (Card et al. 2015), different racial groups (Gerard et al. 2018), and immigrants and natives (Arellano-Bover and San 2020). However, most of this literature does not explore the role of social connections in explaining why some workers find higher-paying jobs than others.[5]

The theoretical literature offers two main mechanisms for the importance of social connections for matching workers and firms. Firstly, social connections might improve the information flow about job opportunities and job seekers (Calvo-Armengol and Jackson 2004; Fontaine 2008). Secondly, connections might impact the value of the prospective match, which may be due to an impact on the productivity of the match (Athey et al. 2000; Bandiera et al. 2009), favoritism (Beaman and Magruder 2012; Dickinson et al. 2018), or to reducing uncertainty about the productivity of the worker or the match (Montgomery 1991; Dustmann et al. 2016; Bolte et al. 2020). In this paper, I build and estimate a matching model that separately identifies these two mechanisms.[6]

Finally, this paper introduces search frictions into a two-sided matching model (Choo and Siow 2006; Galichon and Salanié 2015; Chiappori and Salanié 2016). This extension empowers the model to study labor markets, where search frictions are important (Mortensen and Pissarides 1994; Postel–Vinay and Robin 2002).[7]

2 Data

I use matched employer-employee administrative records from Israel. These data span 1983-2015 and contain administrative information about the entire Israeli workforce collected from tax records. The dataset includes person identifiers, firm identifiers, monthly indicators for each firm in which a person worked, the yearly salary received from each employer in a

ago are not useful.

[5]Two noteworthy exceptions are Schmutte (2015) and Eliason et al. (2019). This paper adds to their contributions by studying how connections matter in an equilibrium model.

[6]Part of this literature also emphasizes the theoretical link between social connections and between-group inequality (Calvo-Armengol and Jackson 2004; Bolte et al. 2020). I use estimates from my model to empirically study this link in the context of the ethnic pay gap in Israel.

[7]The introduction of search frictions to the two-sided matching model also has computational benefits; it enables simulating the model with large-scale data. See Agarwal (2015) for simulation-based estimation of a non-transferable utility model of the market for medical residents. See Jaffe and Weber (2019) for an earlier theoretical study introduces differential meeting rates into Choo and Siow (2006)'s matching model. See also Caldwell and Danieli (2018) for a recent study uses a two-sided matching model to derive a sufficient statistic for studying the effect of outside options on wages.

year, and the firms' industry.

The employment tax records are merged with the Israeli Population Registry. This dataset covers the full population of Israel. It includes demographic information: date of birth, date of death (if any), sex, ethnic group, country of birth, and date of immigration to Israel. Finally, starting in 2000, I observe yearly geocoded information on persons' city of residence.

2.1 Sample selection

I construct a panel dataset at the annual frequency. Following Kramarz and Skans (2014), I assign each person-year observation the firm in which that person was employed during February. I calculate the monthly salary by dividing the yearly salary in a firm by the number of months worked there. If someone worked at more than one firm during February, I assign him to the firm that paid a higher monthly salary. I exclude from the sample worker-year observations with less than 25% of the national average monthly wage.[8] The period of the sample is 1991–2015. I construct a second dataset from this panel dataset, keeping only firms with 5-500 workers per year. I use this data to build a parental network over time.[9]

My analysis sample comprises Israelis who found their first stable job (see definition below) between ages 22-27 in the years 2006-2015 in a 5-500 workers firm. I exclude workers without any parent working in a 5-500 workers firm when they were 12-21 years old. I further exclude immigrants and Ultraorthodox Jews from the sample.

2.2 Definition of first stable job and labor-market entry year

This paper focuses on the employment and salary of young people when they enter the labor market. Following Kramarz and Skans (2014), I define the first stable job as the first job after higher-education graduation (if applicable) that lasts for at least four months

[8]The minimum monthly salary in 2015 was 48.8% of the average salary in that year. This ratio fluctuated between 40%-50% in 1990-2015. Therefore, I exclude workers who earn approximately 50% or less the minimum wage, similarly to Kramarz and Skans (2014). See Appendix B for further details on the data cleaning.

[9]Intuitively, the probability that a random pair of workers form social connections decreases in the firm's size. Below, I show that, indeed, the effect of having a parental connection in a firm on the probability of working at that firm decreases when the firm's size increases. Moreover, I show that the effect disappears for firms with more than 400 workers. Therefore, assuming that a pair of workers in large firms have social connections would increase the error in the measurement of connections, and could downward-bias the estimates of the effect of connections. In 2006-2015, 392 unique firms in Israel employed more than 500 workers (0.2% of the firms). Those firms employ, on average, 37.6% of the labor force. Firms with 1-4 workers account for 70.2% of the firms in that period but employed 10.2% of the labor force. Firms with 5-500 workers, for which this paper studies the effect of social connections, account for 29.6% of the firms in 2006-2015, and employed 52.2% of the labor force (Table A1).

during a calendar year and produces total annual earnings corresponding to at least 150% the national average monthly wage. Labor-market entry year is the year the new worker finds her first job.[10]

2.3 Definition of parental connections

The focus of this study is on the professional network of parents. I study two types of parental professional connections: weak and strong.

Weak connections are connections between children and firms in which precisely one of the parent's past coworkers currently works. Specifically, a child i is (weakly) connected to a firm j if i's parent and a worker k worked simultaneously at the same firm when i was 12-21 years old, and k worked at a firm j at i's labor-market entry year. Both past and current firms should be firms with 5-500 employees.

Strong connections between a child i and a firm j satisfy at least one of the following conditions: 1) i's parent worked at a firm j when i was 12-21 years old, 2) more than one of i's parent's past coworkers worked at a firm j between five years before and five years after i's labor market entry year.

Two components of these definitions are noteworthy. Firstly, to reduce the "endogeneity" in measuring connections, I define the parent's past firms and past coworkers using a fixed period of time (the child is 12-21 years old). I do not include connections that formed at the years between the child is 22 until the year she enters the labor market. Doing so will mechanically increase the set of connections available for workers that enter the labor market later.

Secondly, I assign worker-firm pairs with more than one past parental coworker to the group of strong connections for three reasons. One, it allows me to use the single coworker's characteristics for the classification of the connections. For example, I later define weak and phantom connections by the years the coworker worked at the firm. Likewise, the "death" and "retirement" connections are based on coworker's demographic characteristics. It is

[10] I do not distinguish between the year the fresh graduate looks for her first job and the year she finds her first job. Observing unemployment before starting the first job is difficult in administrative data as only previously employed workers are eligible for unemployment benefits. Potentially, I could use the assignment of workers at some fixed age, or a fixed number of years after graduation, and define people without a job at that time as unemployed. I choose not to do this for two reasons. Firstly, it is challenging to disentangle people who unsuccessfully looked for a job from people who did not look for a job based on employment information alone. For example, many Israeli youths postpone the entry into the labor market because they take a long backpacking trip following military service (Noy and Cohen 2005). Secondly, using the job at a fixed age might bias the estimates of the effect of connections. For example, if the child starts working at the firm before that age and the contact left the firm right after she starts working there, I might define that firm as a firm with phantom connections even though the child had active connections there when she joined the firm.

less clear how to define those concepts when there is more than one contact in the firm. Two, when many parental coworkers work at the same firm, it might be the case that this firm is some continuation of the parent's past firm, e.g., a firm that merged or acquired the parent firm or merely the same firm with a different identifier. Grouping together firms with many parental coworkers with the parents' past firms eliminates weak connections estimates' upward bias. Three, keeping both weak and phantom connections with only one contact makes them comparable, and therefore provides a more accurate estimate for the main effect of interest, namely the effect of weak (indirect) connections. However, I also check the robustness of the results for different definitions of connections (see Table A3).

2.4 Workers' and Firms' characteristics

The paper's empirical analysis compares the firm assignment and wages of new workers with similar observable characteristics. These characteristics include age, gender (male/female), education (no college/college), ethnicity (Jew/Arab), and district of residence.

Firms' characteristics include the industry, location, and firm pay premium for each firm. I use the 3-digit industry code of each firm (2011 Israeli classification). The firms' locations are determined by the median longitude and latitude of the workers' city of residence. The firm pay premiums are estimated using the AKM model (Abowd et al. 1999). These premiums aim to capture the average differences in salary firms pay to similar workers.[11] See Appendix B for further information about the definitions of the variables.

2.5 Summary statistics

Table 1 shows sample sizes and sample means. The new workers' sample—my main analysis sample—includes 220,877 workers, of which 29% are Arabs, 43% are female, and 23% have some college education. The average age at first stable employment is 24, and the average monthly salary is 5,836 NIS (2017 prices).

On average, Jews who enter the labor market earn more at the first job and work at better firms (in terms of pay premiums) compared to Arabs. Additionally, Jews are connected to higher-paying firms via both strong and weak connections. However, the share of workers who find their first job in a connected firm is higher for Arabs than for Jews (Table 1).

[11] The firm premiums are not necessarily a proxy for the productivity of the firms but might capture other factors that lead to differences in salary, such as differential rent sharing. See Card et al. (2018) for a discussion of the AKM model and the critique of it. In this paper's model, I use the AKM firm premiums only to classify firms into bins. The model's "pay premium" of each bin of firms is estimated within the model and not based on the AKM premiums. This alleviates the potential bias of the AKM firm premiums, such as the limited mobility bias or violation of the implied exogenous mobility assumption.

Comparing males and females, males earn more at the first job but work at similar-paying firms to females. Likewise, males are connected to firms with slightly lower rank (in terms of pay premium) compared to females. Finally, the share of workers who find their first job in a connected firm is higher for males than for females.

To better understand the distribution of connections, I group the firms into 5 bins using the pay premiums. Figure 1 shows the number of weak and strong connections within each bin of firms for different groups of workers. Panels A and B show that, on average, Jews and Arabs have the same number of connections with firms at the lowest quintile of pay premiums. However, Jews have more connections with higher-ranked firms than Arabs, and the gap increases as the firm's rank increases. Overall, the quality of connections (in terms of the pay premium of the connected firms) is better for Jews than Arabs.

Females have a slightly higher number of weak and strong connections than males with each of the firm types, except the lowest firm type, where both groups have a similar number of connections (Figure 1, Panels C and D).[12]

3 Empirical framework

3.1 Identification strategy: comparing real and phantom connections

The main question I want to answer in the first part of the paper is how much more likely the average worker is to work in a connected firm than in an unconnected firm. However, a simple comparison between connected and unconnected worker-firm pairs might attribute the effect of omitted variables to the estimated impact of connections. There might be several reasons why a worker is more likely to work in a connected firm, even without connections per se. For example, Galor and Tsiddon (1997) offer a theory claiming that children tend to choose their parents' occupation because of specific human capital transmitted from parents to children. Suppose other workers working at the parent's firm also tend to have this particular human capital. In that case, the child's probability of working at a firm employing one of their parent's previous coworkers might be high because both have the same specific human capital. Another example is geographical proximity that might be correlated with connections and impact the employment probability.

[12]See appendix C for the correlation between the ethnicity and gender pay gaps on the one hand, and firms and measures of the quality of connections on the other. Correlational evidence suggests that, unlike the gender pay gap, most of the ethnic pay gap in Israel is explained by between-firm variation. Likewise, weak and strong parental connections are correlated with higher wages; this correlation accounts for about 20% of the ethnic pay gap.

This paper addresses this potential endogeneity concern by comparing the probability of working in a firm with an active social tie with a firm with "phantom" social connections. Specifically, it compares the likelihood of working in a firm where the employee had active links in the labor-market entry year with the likelihood of working in a firm where the contact had left a short time before or afterwards. Comparing the treatment and the control group indicates whether workers tend to work in firms with connections rather than in otherwise similar firms.

Formally, a child i has a phantom connection at a firm j if i's parent and a worker k worked simultaneously at the same firm when i was 12-21 years old, and k worked at a firm j between five years before and five years after i's labor market entry year, but not in that year. Only firms with 5-500 employees are considered.

3.2 Econometric model: employment

What is a fresh graduate's propensity to work at a firm with social ties relative to a firm without social ties? To answer this question, I compare the probabilities that graduates with similar observable characteristics work at a specific firm. Some of these graduates are connected to the firm, and some are not. Workers' groups include all combinations of ethnicity, gender, education, age, year of first job, and district of residence of the new workers.

Building on Kramarz and Skans (2014), the probability that child i, belonging to group x, starts working in firm j is

$$e_{ixj} = \phi_{xj} + \sum_{c=p,w,s} \delta^c \cdot D^c_{xj} + \epsilon_{ixj}. \tag{1}$$

e_{ixj} is an indicator variable taking the value one if individual i from group x starts working in firm j. ϕ_{xj} is a match-specific effect that captures the propensity that a child from a given group ends up working in a particular firm. D^p_{ij}, D^w_{ij}, and D^s_{ij} are indicator variables capturing whether worker i has phantom, weak, or strong connections to firm j. The parameters of interest that measure the effect of parental connections are δ^p, δ^w, and δ^s. They are an estimate of how much more likely the average firm is to hire a new worker with phantom/weak/strong connections than an unconnected worker from the same group.

Direct estimation of equation 1 is computationally infeasible, as it required one observation per worker-firm pair, which amounts to more than ten billion observations. In order to estimate equation 1, I apply the fixed-effects transformation, proposed by Kramarz and

Thesmar (2013) and Kramarz and Skans (2014).

Let $D_{ij} \equiv \max_c D_{ij}^c, c \in \{p, w, s\}$ be a variable that indicates whether a worker i has any type of connections in firm j. Firstly, I restrict the sample under study to cases in which there is within group-firm variation in D_{ij}. I then aggregate the model by computing, for each group-firm combination, the fraction of workers with connections in the firm who were hired by that particular firm

$$R_{xj}^{CON} = \frac{\sum_{i \in x} e_{ixj} D_{ij}}{\sum_{i \in x} D_{ij}} = \phi_{xj} + \sum_{c=p,w,s} \delta^c \cdot D_{xj}^c + \epsilon_{xj}^{CON} \tag{2}$$

where $D_{xj}^c = \frac{\sum_{i \in x} D_{ij}^c}{\sum_{i \in x} D_{ij}}$ is the share of c-type connections for children in group x who are connected to firm j. Similarly

$$R_{xj}^{-CON} = \frac{\sum_{i \in x} e_{ixj}(1 - D_{ij})}{\sum_{i \in x}(1 - D_{ij})} = \phi_{xj} + \epsilon_{xj}^{-CON} \tag{3}$$

Taking the difference between the two ratios eliminates the firm-group fixed effects ϕ_{xj}

$$R_{xj} \equiv R_{xj}^{CON} - R_{xj}^{-CON} = \sum_{c=p,w,s} \delta^c \cdot D_{xj}^c + \epsilon_{xj}^R. \tag{4}$$

The variable R is computed for each firm-group combination as the fraction of hirees in the firm from the group having any type of connection to that firm minus the fraction of hirees in the firm from the same group having no parental connection to that firm. The right-hand side variables D_{xj}^c , $c \in \{p, w, s\}$ capture the fraction of connected workers from group x who have the specific connection type c to a firm j. The estimates of δ^c from equation (4) measure the effect of the different types of parental connections.[13]

3.3 Event study: Employment probability by the time the links are destroyed

My identification strategy exploits the time the contact of a new worker left her firm relative to the labor-market entry year to compare the probabilities of new workers working

[13] Note that, by definition, $D_{xj}^p + D_{xj}^w + D_{xj}^s = 1$, which means that the independent variables in equation (4) are collinear. However, the estimation of that equation is feasible because the regression is estimated without an intercept.

at firms with and without active connections in that year. To better investigate the timing of the effect, I estimate the time-varying version of equation (1)

$$e_{ixj} = \phi_{xj} + \sum_{c=p,w} \sum_{\tau=-5}^{5} \delta^{c,\tau} \cdot D_{ij}^{c,\tau} + \delta^s D_{ij}^s + \epsilon_{ixj} \tag{5}$$

where $D_{ij}^{c,\tau}$ is a dummy variable which equals one if i has connections of type c at firm j, and the last year i's contact worked at firm j was τ years after i's labor-market entry year. All other variables are defined as before. Note that, for $\tau < 0$, the contact left the firm before time zero (the labor-market entry year), therefore $D_{ij}^{w,\tau} = 0 \; \forall i, j$. Similarly, if i's contact left the firm at time zero, i cannot have phantom connections at that firm: $D_{ij}^{p,\tau=0} = 0 \; \forall i, j$.

This specification compares the probability of worker i working at a firm in which her contact left the firm just before entering the labor market, to the probability of working at a firm in which the contact left the firm shortly after that time. If social connections increase the probability of finding a job at a firm, there should be a non-continuous increase in the estimated effect at time zero.

3.4 Correlation with salary and job tenure: Within-firm estimates

To check the correlation between parental connections and wages, I compare similar workers at the same firm, with and without connections to the firm. To account for factors correlated with parental connections, I compare real and phantom connections. Specifically, the (log) wage of new worker i equals

$$w_i = \sum_{c=p,w,s} \delta^c D_{i,j(i)}^c + \phi_{x(i)} + \psi_{j(i)} + \epsilon_i. \tag{6}$$

where $D_{i,j(i)}^c$ is an indicator variable capturing whether a worker i has connections of type c at her first job, where the possible types of connections are phantom, weak, and strong. $\phi_{x(i)}$ and $\psi_{j(i)}$ are group and firm fixed effects, respectively. As before, the workers' groups include all combinations of ethnicity, gender, education, age, year of first job, and district of residence of the new workers.

Note that this analysis does not identify the causal effect of social connections on wages since it ignores selection. For example, without connections, a hired connected worker may have counterfactually not received an offer at all instead of a different salary. Unlike the matching question where the outcome (working or not) is observed for each worker-firm combination, the outcome of the wage-setting question is only observed if the worker is hired by the firm. The model addresses this issue by jointly studying questions of matching and wage-setting.

I also check the relationship between parental connections and job tenure. To do so, I run the same regression with the number of years at the first job as an outcome variable.

4 Regression results

4.1 Employment

This section estimates the probability that the worker finds her first stable job in a firm where she has parental connections using equation (4). The connection effects (δ_c) capture the excess probability of a new worker finding her first job at a c-connected firm compared to a worker without any connections. I simultaneously estimate the effect for the three types of parental connections defined above, namely phantom, weak and strong connections. Comparing the impact of weak and strong connections, and phantom connections, allows me to isolate the effect of connections from other factors that might be correlated with them.

Even after the fixed-effects transformation, limited computational resources prevented estimation of the model using all observations. Therefore, I take a 20 percent random sample of the new workers in each iteration and run 100 such iterations. Using the distribution of estimates obtained, I calculate the mean and 95 percent confidence intervals of the regression coefficients and the other statistics of interest.

Table 2 presents estimates of the coefficients in equation (4). Each column shows a separate estimate for a different population group based on ethnicity and gender. All estimates of the effect of the three types of connections are positive and statistically significant, which implies that new workers with any connections to a firm are more likely to work there than workers with similar observable characteristics but no connections to the firm.

The regression results show that the effect is much more substantial for weak and strong connections relative to phantom connections. Having weak (strong) connections at the firm increases the probability of working there by 0.05 (0.49) percentage points relative to someone with no connections. In contrast, phantom connections increase this probability only by 0.01 percentage points. To better understand the magnitude of the effect, I calculate the ratio between the likelihood of working in weakly- or strongly-connected firms and phantom firms. The estimated probability of working in a weakly (strongly)-connected firm was 3.7 (32.5) times higher than the probability of working in a phantom-connected firm for the average new worker (Table 2, column 1).

Columns 2 and 3 of Table 2 report the estimated effects separately for Jews and Arabs, the two main Israeli ethnic groups. The estimated impact of weak connections was stronger for Arabs than for Jews; the probability of working in a weak-connected firm was 4.2 times higher than a phantom-connected firm for Arabs, and 3.3 times for Jews. Similarly, the effect of weak connections was stronger for males (4.4) than females (2.7) and was stronger for new workers with no college education (3.9) than for new workers with at least one year of college (3.0, Table 2, columns 4-7).

4.2 Event study

The estimates of the coefficients in equation (5) are plotted in Figure 2—the probability of working in a firm as a function of the lag between the last year the potential contact worked at the firm and the labor-market entry year. Negative lags represent phantom connections, and non-negative lags represent weak connections.[14]

The probability that a new worker began work at a firm that her parental contact left before she entered the labor market was higher by 0.005-0.012 percentage points than the probability of another worker with similar observable characteristics but no connections at all. The estimated effect increased to 0.040-0.057 percentage points when the contact left the firm after time zero. The discrete increase in the employment probability happens exactly at time zero—the labor-market entry year, indicating that the existence of the contact at the firm at that time accounts for the change in the probability of employment.

4.3 Balancing test

My identification strategy assumes that, without parental connections, there is no systematic difference between the probability of working in a firm with a weak (active) connection and in a firm with a phantom (non-active) connection. I support this assumption in three ways.Firstly, I show that firms with weak and phantom connections are similar on a variety of characteristics such as sector and location. Secondly, I estimate the effects using two exogenous causes of separation, coworkers' deaths and retirements, and show that the estimates obtained are similar in magnitude to the benchmark result using all causes of separation.

[14] The figure does not show the estimates for strong connections and phantom connections in which the potential contact left the firm after time zero but did not work there at time zero (for example, she started to work at the firm after that time). Table A4 reports all estimated coefficients of equation (5). The estimated effect for strong connections is of a similar magnitude to that in the benchmark model presented in Table 2. The estimated effects for phantom connections with positive lag is significantly smaller than the parallel effect for weak connections.

Finally, I perform a placebo test, assigning a worker's connections to a random worker with similar observable characteristics, and find no hiring differences between phantom and real connections of a random worker.

I start with the balancing test. As mentioned earlier, social connections between a worker and a firm might be correlated with other similarity measures between the worker and the firm. Two leading examples are the geographical distance between the worker and the firm and the similarity between the firm and the firms in which the worker's parents have worked. Indeed, in what follows, I show that the distance between workers and firms is smaller if there are parental connections between the worker and the firm. Likewise, the probability that the firm is in the same industry as one of the parent's firms is higher if there are connections. In the first test of the identification strategy, I check whether there are also such differences between phantom and real parental connections.

To do so, I re-estimate equation (1) with the distance/similarity measures as the outcome variable. The first measure is the distance between the worker's location at age 21 and the firm's location.[15] Column 1 of Table A2 shows the estimated coefficients. As expected, compared to firms with no connections, firms with all three types of social connections are significantly closer to the workers' locations. However, the estimates for phantom and weak connections are virtually identical, with -0.369 and -0.368 log points.

The second measure is an indicator variable that equals one if the firm has the same 3-digit industry code as one of the parents' previous firms. Once again, connected new workers were more likely to have parents who worked in the same industry, compared to unconnected workers. This correlation, however, is similar to phantom and weak connections, with estimates of 0.077 and 0.076 percentage points, respectively (Table A2, column 2).

4.4 Exogenous separation: Death and retirement of potential contacts

This paper's identification strategy exploits the timing of workers' parents' coworkers employment relative to the workers' labor market entry. I assume that other than the effect of social connections at the time of the job search, there should be no systematic difference in the probability of working in real- and phantom-connected firms. This assumption breaks if the separation time is correlated with other factors, unrelated to social connections, that affect employment decisions. For example, workers that leave a firm might deliver to their contacts a negative opinion about the possibility of working at that firm. This mechanism

[15] I do not use the worker's location at the labor-market entry year to avoid the mechanical correlation between the workers' locations and the firm as a result of moving closer to the workplace.

would decrease the probability of working at the firm only for workers whose contacts left the firm before they started to work, not after. In this case, having phantom connections at the firm would decrease the job seeker's probability of working there compared to real connections.

To further investigate this possibility, I estimate the effect of connections for two exogenous reasons for separations. The first specific separation cause is death. I classify the separation cause as "death" if the contact died not more than one year after working at the firm. I compare the probability of working at firms where the (dead) potential contact worked at the firm before time zero to the probability of working at firms in which the connection worked at the firm after that time and died immediately afterward.

The second separation cause is quitting the job precisely at retirement age. In Israel, the statutory retirement age is 62 for females and 67 for males. At that age, workers are entitled to leave their job and receive a pension. Figure A2 plots the distribution of workers' ages in the last year of employment for males and females. This figure shows that it is common to leave the labor force at the retirement age. I compare workers that quit their firm at the retirement age, before and after year zero.

For each special type of connection, I split the set of phantom and weak connections into two subsets, each with connections belonging to the death/retirement group (i.e., the contact died or left the job at the retirement age), and connections that do not belong to that group. I then re-estimate equation (1) using the five types of connections (phantom-death/retirement, phantom-other, weak-death/retirement, weak-other, and strong).

Table 3 reports the results of this exercise. Compared to fresh graduates without connections to the firm, the probability of working at the firm with a contact that died while employed at the firm or immediately afterward is higher by 0.031 percentage points if the last year the contact worked at the firm was before time zero and by 0.065 percentage points if it was after time zero. The estimates for firms with other contacts, i.e., contacts who did not die at the year after leaving the firm, are virtually identical to the baseline results (0.01, 0.05, and 0.49 for phantom, weak, and strong connections, receptively). The ratio between the probability of working in a firm with weak connections compared to a firm with phantom connections is 2.6 for "dead" connections and 3.7 for other connections (Table 3 column 1). However, due to the small number of such cases, the estimated ratio for "dead" connections is not significantly different from 1.

Similar results were obtained when using the statutory retirement age as a special case of job separation. Once again, the estimates for firms with contacts who left the firm exactly at their retirement age is higher for weak connections than phantom connections (0.01 and 0.03 percentage points, respectively). The ratio between weak and phantom connections is 3.9 for connections in firms where the contacts that left the firm at the retirement age, compared to 3.7 for other connections. I also estimate the effect by combining the death and retirement causes of separation. The estimated ratio between weak and phantom connections is 2.8, compared to 3.7 for other connections. These ratios are not statistically significantly different from 1 (Table 3 columns 2 and 3).

Overall, the estimated effects of connections are quantitatively similar for contacts who left the firm for "exogenous" reasons (death or retirement) and other contacts. The ratio between the probability of working in a firm with weak connections and in a firm with phantom connections is slightly smaller for "death" and somewhat larger for "retirement," compared to other connections. However, due to the relatively small number of connections belonging to these types, the estimates of the special types of connections are much noisier. These results suggest that the estimated effects of connections obtained from the benchmark model (with all connections) are not a result of endogenous separation that differentially impacts phantom and weak connections, but the effects of the connections themselves.

4.5 Placebo test: Assigning worker's connections to another worker

Another threat to the identification strategy is if firms with different types of connections have different hiring trends. For example, one might think that firms with more phantom connections with new workers tend to be on a downward trend in employment and are therefore hiring fewer new workers regardless of the impact of connections.

To address this concern, I perform a placebo test and assign a worker's connections to another worker in her group. If the employment probability gap between real- and phantom-connected firms is mediated by other factors correlated with these types of connections, the probability of a worker working in a firm that another member of the group has real connections to should be higher than in a firm with phantom connections.

Table 4 reports the estimates of equation (1) assuming each worker has the set of connections of a random member of her group. None of the estimates are statistically significantly different from zero. Moreover, there is no statistically significant difference between the estimated probability of working in a weak-connected firm relative to a phantom connected firm. The event study estimates of equation (5) also showed no difference between phantom and real connected firms (Figure 3).

4.6 Robustness check: changing the definitions of parental connections

In the baseline specification, I combined firms with direct connections (parents' past firms) and firms where multiple of the parents' past coworkers worked later, in the group of "strong connections".[16] The first column of Table A3 reports the baseline specification again, where direct connections and multiple indirect connections (either real, phantom, or any combination of them) are grouped. In the second column, I estimate a separate coefficient for direct and multiple contacts. The coefficients of weak and phantom connections are 0.012 and 0.053, almost identical to the benchmark model with estimates of 0.010 and 0.050, respectively. The ratio between the probability of working in a weakly connected firm compared to a phantom connected firm is 3.4, compared to 3.7 in the benchmark model. The estimated coefficients for direct and multiple contacts are 3.091 and 0.171; both are significantly greater than the coefficient of weak connections. Comparing to the baseline model, the effect of strong connections, which combined direct and multiple connections, is 0.487, lower than the estimate for direct connections alone and higher than that for multiple connections alone.

In the third column of Table A3, I combine single and multiple phantom connections into one group. Likewise, I combine single and multiple weak connections into one group. If both phantom and weak connections work at one firm, I assign than firm to the group of weak connections. The coefficients for phantom and weak connections are now 0.015 and 0.095, respectively, greater than the estimates from the benchmark model. The estimate for the effect of direct connections is now 3.092. The weak-phantom ratio is 5, greater than the ratio in the baseline model.

Taken together, the results of this robustness check indicate that the estimated effects using the baseline definition of parental connections can be seen as a lower bound for both the effects of indirect and direct connections. The impact of multiple contacts in a firm on the employment probability is stronger than the effect of indirect connections but weaker than direct connections. When combining single and multiple indirect and phantom connections in the same group, the effects of both weak (indirect) and strong (direct) connections is larger.

4.7 Heterogeneity of the effect

Is the impact of parental connections on employment similar for workers who belong to different groups? How do the characteristics of the connections themselves change the effect?

[16] See the discussion in section 2.3.

To check the heterogeneity of the effect, I re-estimate equation (1) with separate coefficients for different groups of weak connections. Figure 4 shows the results. Below are the main findings.

Past and current firms' size. Connections that formed at smaller firms are more effective (Figure 4, Panel A). The effect disappears for firms with more than 400 workers. This result is consistent with the intuitive view that the probability/intensity of the connections between a random pair of workers is higher the smaller is the firm. Moreover, finding a job in a connected firm is more likely in smaller firms (Panel F). This fact also can be explained by a higher probability that the contact can impact the hiring decisions in smaller firms.

Parent's and coworker's salary rank. I check both the countrywide salary rank and the rank within the firm. Panels B and D of Figure 4 show that, except for the very low wage percentiles, the overall relationship is negative, indicating that workers from high-income background use connections less. This result is correct both concerning the wage rank of the parent and the coworker. On the contrary, the firm's salary rank is positively correlated with the effect (Figure 4, Panels C, and E).

Length of co-working and time since co-working. As expected, the effect is more substantial the longer the parent and the contact worked together. The effect is weaker for connections generated less recently (Figure 4, Panels H and I).

Gender. The effect is stronger for males than for females. This fact is true for the gender of the worker, the parent, and the parent's coworker (Figure 4, Panels J-I).

Ethnicity and education. The effect is stronger for Arabs than Jews and weaker for more highly-educated workers (Panels M-O). This result is consistent with the findings above that the effect is stronger for workers from a lower socio-economic background.

Similarity between the child, the parent, and the coworker. The effect is stronger if the parent, the worker, or the parent's coworker are the same gender. Likewise, the effect is stronger if the worker and the parent's coworker are from the same ethnic group. Finally, the smaller the wage gap between the parent and the coworker, the stronger the effect (Panels G, P, Q, and R).

4.8 Correlation with salary and job tenure

So far, I found that parental connections in a firm increase a worker's probability of having her first job there. Next, I turn to check the relationships between parental links and other labor-market outcomes of new workers. This subsection compares the wage and job tenure of connected and unconnected workers working at the same firm.

The first column of Table 5 reports the estimates of equation (6) with log salary as the outcome variable. The salary of workers with phantom connections is higher by 1.2 log points than observably similar workers at the same firm without connections. On the other hand, having real connections at the firm, either weak or strong, is correlated with a higher salary than workers without connections. The coefficients are 2.6 and 8.2 log points for weak and strong connections, receptively. Compared to phantom connections, weakly and strongly connected workers' salaries were higher by 1.4 and 7.1 log points.

The second column of Table 5 investigates whether workers with a connection at their first firm stay at that firm for more extended periods than unconnected workers. The outcome variable in column 2 is the number of years the worker stayed at her first firm. The first-job duration of workers with phantom, weak, and strong connections is higher by 0.098, 0.187, and 0.441 years, respectively, compared to workers without connections. Compared to phantom links, weak and strong connections are correlated with 0.089 and 0.343 more years at their first firm.

Overall, this subsection shows that, on average, connected workers receive higher wages at the firm and stay at the same firm for longer periods. Comparing worker-firm pairs with real and phantom connections helps isolate the relationships between these outcomes and social relations from other factors correlated with connections, such as geographical distance and industrial similarity. However, reduced-form estimates cannot separate the effects of connections on the firm's salary from the worker and firm joint match decision. The structural model in the next section addresses this issue by jointly studying questions of matching and wage-setting. The wage differentials between connected and unconnected workers are translated into differences in the expected firm's utility for different worker-firm matches. Likewise, although not explicitly modeled, the correlation between parental connections and job duration is consistent with higher match utility the firms get from hiring connected workers. I discuss these issues in more detail below.

5 A Matching Model of the Labor Market

In the second part of the study, I build and estimate a two-sided matching model of the labor market. Typically, this literature assumes that each agent has perfect information about all other agents in the economy (Choo and Siow 2006; Galichon and Salanié 2015; Chiappori and Salanié 2016). Agents choose a pairwise stable match in which there is no pair of unmatched workers and firms that strictly prefer each other. In my model, I restrict the feasible choice set to consist of only pairs that have previously met.[17] With this

[17] Del Boca and Flinn (2014) also study a matching model (of the marriage market) with restricted choice sets.

assumption, I depart from the perfect information assumption by adding "search frictions" into the matching model.

Matching takes place in two stages. In the first stage, workers and firms meet randomly, and the probability of meeting depends on the observable characteristics of the worker, the firm, and the connections between them. In the second stage, workers and firms that have met choose their optimal (stable) match, based on the utility they obtain from the match.

Using this conceptual framework, I separate the potential mechanisms of the effects of connections on firm assignment and wages into two groups. The first mechanism is that social connections reduce the job search frictions by improving the information flow about open vacancies and potential candidates. The second mechanism is that connections directly impact the value of the match. This effect might be due to a direct impact on the match productivity, favoritism, or better information about the worker's characteristics, or the prospective match quality.

The purpose of the model is twofold. First, it allows the evaluation of counterfactuals accounting for equilibrium effects. For example, generating new connections between a set of workers and a set of firms might directly decrease the probability of other workers to work in those firms or affect the structure of wages in the economy.

The second role of the model is to disentangle the two mechanisms described above. This question is essential to predict the effectiveness of different policy measures. For example, suppose connections are useful mainly because of they alleviate "search frictions". In that case, one might think that policies that aim to create job interviews between workers from disadvantaged groups and good firms can substitute social links.[18] However, it is less plausible to assume that such policies can generate additional value to the match. Therefore, if the "match utility" is dominant, the effectiveness of such policies will be more moderate.

5.1 Setup

Each worker i belongs to one observable group $x \in \mathcal{X}$ in a market $t \in \mathcal{T}$. Likewise, each firm j belongs to one observable group $y \in \mathcal{Y}$ in a market $t \in \mathcal{T}$. There are I_{tx} workers of type x in market t, and J_{ty} firms of type y in market t. In each market t, the overall number of workers, I_t, and the overall number of firms, J_t are equal. Each firm/job belongs to a specific year and can employ only one worker. Much like most of the matching literature, the model is static. Each worker i and firm j are connected by exactly one type of connections $c = 0, 1, ..., C$. In practice, I use the same three types of connections as above, namely phantom, weak and strong connections. $c = 0$ denotes no connections.

[18] An example of such policy is "The Rooney Rule," which requires NFL teams to interview at least one minority candidate any time their head coaching position comes open (Solow et al. 2011).

The matching process takes place in two stages. In the first stage, workers and firms randomly meet. Let m_{ij} be a binary variable equal to one if there is a meeting between worker i and firm j, then

$$m_{ij} = 1\left(\rho_{ij} \leq p_{ij}\right) \tag{7}$$

where ρ_{ij} is a draw from an i.i.d. standard uniform distribution, and p_{ij} is the meeting probability based on the observable characteristics of i and j. Only workers and firms from the same market can meet. Finally, denote $m_i = \{j|m(i,j) = 1\}$ and $m_j = \{i|m(i,j) = 1\}$.

In the second stage, there is a matching process between all workers and firms in each market, with the restriction that workers and firms that did not have a meeting at the first stage cannot form a match. Following Choo and Siow (2006), I assume transferable utilities (TU). The utility of a firm j which employs a worker i is

$$V_{ij} = f_{ij} - w_{ij} \tag{8}$$

where f_{ij} is the firm's utility from the match, and w_{ij} is the wage the firm pays to the worker. The utility of workers is simply the wage they get

$$U_{ij} = w_{ij}. \tag{9}$$

5.2 Equilibrium

I follow the matching literature and use the pairwise stable matching for the definition of equilibrium.

Definition 1 (equilibrium outcome). An equilibrium outcome (μ, w) consists of an equilibrium matching $\mu(i,j) \in \{0,1\}$ and an equilibrium wage $w(i,j) \in \mathbb{R}$ such that

1. Matching $\mu(i,j)$ is feasible

$$\sum_{j=1}^{J} \mu(i,j) \leq 1 \quad , \quad \sum_{i=1}^{I} \mu(i,j) \leq 1 \quad , \quad \mu(i,j) = 1 \implies m(i,j) = 1 \tag{10}$$

2. Matching $\mu(i,j)$ is optimal for workers and firms given wages w and meetings m

$$\mu(i,j) = 1 \implies j \in \text{argmax}_{j \in m_i} U_{ij} \quad \text{and} \quad i \in \text{argmax}_{i \in m_j} V_{ij} \tag{11}$$

5.3 Finding the equilibrium matching

Let $\pi_{ij} = U_{ij} + V_{ij}$ be the joint surplus from a match between worker i and firm j. Shapley and Shubik (1971) show that μ is an equilibrium matching if and only if it maximizes the total joint surplus

$$\begin{aligned} \mu \in \text{argmax}_{\mu} \sum_{\mu(i,j)=1} \pi_{ij} \\ \text{s.t.} \quad \mu \text{ is feasible, i.e., equation (10) holds} \end{aligned} \tag{12}$$

This claim transforms the decentralized matching problem into a centralized assignment problem. To find the equilibrium matching, we need to find the assignment that maximizes the total surplus. The assignment problem can be solved by linear programming or auction algorithms. In practice, I find the auction algorithm much faster for the problem at hand. The description of the auction algorithm here follows Bertsekas (1998).

Definition 2 (the auction algorithm)

1. Start with an empty assignment S, a vector of initial payoffs u_i, and some $\epsilon > 0$
2. Iterate on the two following phases
   (a) Bidding Phase
   Let $J(S)$ be a nonempty subset of firms j that are unassigned under the assignment S. For each firm $j \in J(S)$
      i. Find a "best" worker i_j having maximum value, i.e.,

$$i_j = \arg\max_{i \in m(j)} \pi_{ij} - u_i \tag{13}$$

      and the corresponding value

$$v_j = \max_{i \in m(j)} \pi_{ij} - u_i \tag{14}$$

and find the best value offered by workers other than i_j

$$q_j = \max_{i \in m(j), i \neq i_j} \pi_{ij} - u_i \tag{15}$$

ii. Compute the "bid" of firm j for worker i given by

$$b_{ij} = u_{i_j} + v_j - q_j + \epsilon \tag{16}$$

(b) Assignment Phase

For each worker i, let $B(i)$ be the set of firms from which i received a bid. If $B(i)$ is non-empty, increase u_i to the highest bid

$$u_i = \max_{j \in B(i)} b_{ij} \tag{17}$$

and assign i to the firm in $B(i)$ attaining the maximum above

3. Terminate when all workers are assigned to firms

Bertsekas (1998) showed that if at least one feasible assignment exists, the auction algorithm terminates with a feasible assignment within $I_t \cdot \epsilon$ of being optimal, where I_t is the number of workers (and firms) in the market. Moreover, there exists a small enough ϵ such that the auction algorithm terminates with the optimal assignment.

The auction algorithm's practical performance is considerably improved by applying the algorithm several times, starting with a large value of ϵ and successively reducing it up to some final value $\hat{\epsilon}$ such that $I_t\hat{\epsilon}$ is deemed sufficiently small. Each application of the algorithm provides good initial payoffs for the next application (Bertsekas 1998). In practice, I exploit the data's sparsity using the implementation of the auction algorithm proposed by Bernard et al. (2016).

The proposed two-stage model offers a computational advantage over existing matching models. If M is the average number of meetings per worker, then in each market, there are about $(M-1)!I_t$ possible allocations, relative to $I_t!$ in the unconstrained matching problem. That means that the optimal allocation can be found for small enough M, whereas it cannot be found in standard matching models for large datasets. This computational advantage allows the estimation of a matching model based on simulations, which allows a richer set of specifications for the systematic and idiosyncratic utilities in the model.

5.4 Finding the equilibrium payoffs

Generally, if there exists a feasible matching, there exists a unique equilibrium matching (Shapley and Shubik 1971).[19] However, the equilibrium payoffs that support the equilibrium matching are not unique. First, note that if w is an equilibrium wage schedule, so is $w+r$.[20] Therefore, one needs to normalize the location of wages in each market.[21]

Second, even after that normalization, the set of equilibrium payoffs is generically not a singleton. Let u_i be the the payoff of worker i in equilibrium. Demange and Gale (1985) show that u_i is a lattice. That is, there exist $\{\underline{u}_i, \bar{u}_i\}_{i=1}^{I}$ such that $\{u_i | \underline{u}_i \leq u_i \leq \bar{u}_i\}_{i=1}^{I}$ is the set of equilibrium payoffs.

In words, the set of equilibrium payoffs is characterized by component-wise upper- and lower- bound payoffs. The upper bound payoffs correspond to the workers' preferred equilibrium, while the lower bound payoffs correspond to the firms' preferred equilibrium (Bonnet et al. 2018).

In a standard matching model, when every worker can be matched with every firm, the set of equilibrium payoffs shrinks to a singleton when the number of agents goes to infinity (Gretsky et al. 1999). This result is not true in the current model, in which the meeting requirement restricts matching. In this case, the set of equilibrium payoffs shrinks to a singleton only when the number of meetings per worker goes infinity. In practice, I simulate the model with a relatively small number of meetings per worker; therefore, the set of equilibrium payoffs has a non-trivial range that has to be found.

Given the equilibrium matching, the bounds on the equilibrium payoffs can be found using the Bellman-Ford algorithm (Ahyja et al. 1993; Bonnet et al. 2018). Let u_i and v_j be the equilibrium payoffs for workers and firms, respectively, in a connected set G, with the normalization $u_1 = 0$. The firm-optimal equilibrium payoffs are the fixed point of the

[19] This is true under standard regular conditions. For example, if the joint surplus π_{ij} is coming from a continuous distribution, then with probability one, the equilibrium matching is unique.

[20] I assume that I do not observe unmatched workers and firms ("singles") in the data. Therefore the model does not include outside options that might restrict the payoff location. See Dupuy and Galichon (2014) for the case that singles are observed.

[21] Formally, consider the set of meetings between workers and firms as a non-directed graph G. A market is a connected subgraph of G.

mapping

$$u_i = \max(u_i, \max_{j\in m(i)} (\pi_{ij} - v_j)) \quad , \quad v_j = \min(v_j, \pi_{i^*(j)j} - u_{i^*(j)}) \quad , \quad u_1 = 0 \tag{18}$$

where $i^*(j)$ denotes the equilibrium match of firm j. The fixed point of this map can be computed by iterating on (18) from the initial values $\{u_i = -\infty, u_1 = 0; v_j = \infty\}$. Similarly, the worker-optimal equilibrium payoffs can be computed by iterating on

$$v_j = \max(v_j, \max_{i\in m(j)} (\pi_{ij} - u_i)) \quad , \quad u_i = \min(u_i, \pi_{ij(i)} - v_{j(i)}) \quad , \quad u_1 = 0 \tag{19}$$

from the initial values $\{u_i = \infty, u_1 = 0; v_j = -\infty\}$.

Definition 3 (Double-connected set). A double-connected set of nodes is a connected set in which each node is connected to at least two other nodes.

Claim 1 (existence of finite wage bounds). Let G be the graph of meetings. Let $\{G_1, ..., G_T\}$ be the set of connected subgraphs of G. Assume that in each subgraph G_t, the number of workers and firms is equal and normalize the first worker's payoffs in each subgraph to zero. Then, the finite upper- and lower- bounds $\{\underline{u}_i, \bar{u}_i\}_{i=1}^{I}$ are finite if and only if all subgraphs are double connected.

Proof. Let G_t be a double connected set. Let $\{\underline{u}_i\}_{i=1}^{I_t}, \{\bar{v}_j\}_{j=1}^{J_t}$ be the firm optimal payoffs. Assume by contradiction that $\underline{u}_i = -\infty$ for some $i \in \{2, ..., I_t\}$. Because G_t is double connected, there exists a firm $j \neq j^*(i)$ belonging to $m(i)$. Let v_j be an equilibrium payoff of j. Because $\underline{u}_i = -\infty$, there exist small enough u_j such that $u_j < \pi_{ij} - v_j$. But this contradicts the optimality of the match. The symmetric argument holds for the worker optimal payoffs.

Now, assume G_t is not double connected. WLOG, assume there exists a worker such that $|m(i')| = 1|$. Assume by contradiction that $\underline{u}_{i'} > -\infty$ is finite. Let $\{u_i\}_{i=1}^{I}, \{v_j\}_{j=1}^{J}$ be equilibrium payoffs, and note that changing only the payoff of i' to $u_{i'} = \underline{u}_{i'} - 1$ still supports the same equilibrium matching. ■

To avoid the pathological cases of nodes with less than two edges, I assign two extra meetings for each worker and firm in each simulation, regardless of the meetings they draw based on the parameters. Precisely, let $i = 1, ..., I_t$ be the sequential number of workers and firms in market t. I draw two random permutations of length I_t, Per^1 and Per^2, such that

$Per^1(i) \neq Per^2(i) \quad \forall i = 1, ..., I_t$, and assume that worker i has meetings with firms $Per^1(i)$ and $Per^2(i)$.[22]

To get a unique prediction of the equilibrium wages, I assume the payoffs are

$$u_i = \lambda \underline{u}_i + (1 - \lambda)\bar{u}_i \tag{20}$$

for some $\lambda \in [0, 1]$. In practice, I simply assume $\lambda = 1/2$.

5.5 Parametrization

I assume a flexible model in which the meeting and utility parameters are potentially different for each combination of market t, worker group x, firm group y, and connection type c. Specifically, the meeting probability between worker i and firm j depends on their observable characteristics

$$p_{ij} = p_{txyc}. \tag{21}$$

Likewise, the utility of a firm j is

$$log(f_{ij}) = b + \beta_{txcy} + \sigma \cdot \xi_{ij} \tag{22}$$

where β_{txcy} is the systematic utility and depends on the observable characteristics of i and j, and ξ_{ij} is drawn from a standard normal distribution. σ is a parameter that needs to be estimated. In line with the AKM literature that assumes the log wage is additive in worker's and firm's effects, I assume a log-linear specification of the firm's utilities, which are closely related to the wages.

The meeting probability and the firm's systematic utility depend on the year, worker characteristics, firm characteristics, and connection characteristics. In the estimation, I assume that each year is a separate job market and consider the new workers from my sample who find their first job in that year and the jobs that have been found as the participants of the matching game that year. As in the reduced-form part, the years are 2006-2015 (ten

[22] As these extra meetings are orthogonal to the model's parameters, there is no impact on the estimated parameters. One obvious exception is the meeting parameters' level, which needs to be increased by an average of two meetings per worker. However, as explained below, that level is not identified in the current model, and is normalized to a fixed level.

years). To classify workers, I use three binary characteristics: ethnicity (Jew/Arab), gender (male/female), and education (no college/some college or more). I classify workers into eight groups based on all the possible combinations of these characteristics. Likewise, I classify firms into five bins of AKM pay premium. There are four categories of connections between a worker and a firm: no connections, phantom connections, weak connections, and strong connections. Overall, there are $10 \times 8 \times 5 \times 4 = 1,600$ cells of observable characteristics.

5.6 Moments

There are three sets of parameters in the model that need to be estimated: the firm's systematic utilities β_{txyc}, the meeting probabilities p_{txyc}, and the idiosyncratic standard deviation σ. To estimate them, I use three sets of moments obtained from the data. The first is the number of matches in each (t, x, y, c) cell μ_{txyc}. The second is the average wage in each cell w_{txyc}. The last moment is the wage variance Var_w. Denote the set of all moments by $h = (\mu_{txyc}, w_{txyc}, Var_w)$.

In practice, I divide each firm into several one-worker firms (or jobs) each year according to the number of new matches observed in the data. However, to determine the connection type between a firm/job and a worker, I use the definitions from the first part of the paper. Thus, if a firm hires multiple workers in one year and a worker i has connections of type c to that firm, I assume that the worker has a connection of type c to each of the firms/jobs belonging to the original firm. See Appendix D for further information on the calculation of the moments.

Under the parametric assumptions described above, for a given parameter vector $\theta = (\beta, p, \sigma)$ and a draw of the unobservables $\zeta = (\rho, \xi)$, a unique equilibrium matching $\mu_{ij}(\theta; \zeta)$ and wages $w_{ij}(\theta; \zeta)$ exist and can be simulated:

1. Get the set of meetings m_{ij}
2. Calculate the joint surplus π_{ij}
3. Find a feasible matching that maximizes the total joint surplus using the auction algorithm
4. Find the wage's bounds using the Bellman-Ford algorithm

Using the equilibrium outcome, I can compute the model analogous to the data moments $\hat{h}(\theta; \zeta) = (\hat{\mu}_{txyc}(\theta; \zeta), \hat{w}_{txyc}(\theta; \zeta), \hat{Var}_w(\theta; \zeta))$.

5.7 Identification

This section discusses, informally, some of the identification issues of the model. Assume that $\hat{h}(\theta^1, \zeta) = h$ for some θ^1 and ζ. Identification requires that $\hat{h}(\theta^2, \zeta) \neq h$ for every $\theta^2 \neq \theta^1$. Firstly, assume that p and σ are known and only β is unknown. This model is similar to standard matching models, and data on matches alone is enough for the identification of β (Galichon and Salanié 2015; Salanié 2015).

Secondly, assume that only σ is known. In this case, using the information on matching only without wage data, we cannot separately identify the two underlying parameters of the model, namely the meeting probabilities and the match utilities. A high number of matches of a group of workers and firms could happen because the group's meeting rate is high or because the utility of those matches is high. However, the two parameters can be separately identified using both matching and wage data. The reason for that is that the two sets of moments, namely the groups' number of matches and wages, react differently to changes in the meeting rate and utility parameters. The group's match value is important both for the groups' number of matches and wages. In contrast, the group's meeting rate is important much more for matches than for wages.

To see the intuition for that, consider a single worker i and assume that she draws M iid wage offers from some distribution from firms in each of Y bins. Assume that the worker is choosing to work at the firm offering the highest wage. Now, consider two interventions: 1) Increasing the value of each draw of firms of type y by t percent. 2) Increasing the number of draws from firms of that type by t percent. In the first intervention, the impact on both the worker's probability of working at a firm of type y and the expected wage is large. In contrast, in the second innervation, only the impact on the probability of working at a firm of type y is large, but the impact on the expected wage is moderate and goes to zero as MY is getting large. The same intuition holds when considering equilibrium effects.

To check if the model predictions fit the intuitive arguments mentioned, I run 10,000 simulations. Each time, I change the value of only one parameter of one xyc group in each market t. Then, I compute the difference between the model's moments with the new and old parameters.

Figure 5 plots the distribution of the moment differences for the same $txyc$ group of workers and firms for which the parameter is changed. As expected, a positive shock to the meeting probability and the group's utility positively impact the number of matches for that group predicted by the model (Panels A-B). Also, there is a positive change to the group's average wages, given a change in the utility parameter (Panel C). However, a change in the meeting parameter has very little impact on wages (Panel D).

Table A5 reports the simulated elasticities between the moments and the model's parameters. The first row shows the same group of workers and firms for which the parameter is changed. The matches-utility, matches-meetings, and wages-utility elasticities are all positive and large, with estimated values of 3.51, 0.77, and 3.43. However, the wages-meetings elasticity is only 0.015, which is of the same order of magnitude as the indirect effects reported in the second row of Table A5. This small increase is due to a better choice set for the workers.

Now, assume θ^2 is identical to θ^1 except for the meeting and utility parameters of one $txyc$ group. Assume by contradiction that $\hat{h}(\theta^2, \zeta) \neq h$. If only one of the parameters is different, then because of the monotonicity of μ_{txyc} with respect to both p_{txyc} and β_{txcy}, we have $\hat{\mu}^1_{txyc} \neq \hat{\mu}^2_{txyc}$. Next, assume WLOG that $\beta^2_{txcy} > \beta^1_{txcy}$. Because the number of matches is increasing in β, it must be the case that $p^2_{txcy} < p^1_{txcy}$. But because the wages are (almost) not impacted by p, this implies $w^2_{txcy} > w^1_{txcy}$.[23]

Thirdly, identification of σ comes, again, from the fact that we observe wages. If wages are not observed, only the ratio between the match utility and the idiosyncratic utility is identified using matching information. However, when wages are also observed, both the scale of the match systematic value and the amount of unobserved heterogeneity necessary to rationalize the data can be identified (Dupuy and Galichon 2015). I use the variance of the wages to pin down σ.

Finally, the level of p_{txyc} is not identified together with the other parameters of the model. In a standard matching model (without the meeting restriction), the unobserved heterogeneity is the only source of imperfect sorting on observable characteristics. The meeting restriction adds another channel for the imperfect sorting: even if some pairs want to match if they knew each other, they cannot do so because of the information friction. But these two channels cannot be separately identified based on the observed amount of sorting. To see why, assume that we double the number of meetings per worker for all groups. That would result in a better (observable) sorting. But that could also be done by decreasing the amount of unobserved heterogeneity in the model. In the estimation, I normalize the first

[23] I did not show the identification in the case that the parameters of more than one $txyc$ group are different. The intuition is that the direct effect of changing the parameters of one $txyc$ group on the matches and wages of the same group is much stronger than the indirect effect of another group's parameters, say $txy'c'$, on the moments of $txyc$. Then, we need a larger change to the parameters of $txy'c'$ such that the indirect effect is equal to the direct effect. But then the moments of $txy'c'$ are different from the true moments. This argument can be extended to more than two groups. A formal proof of this argument is beyond the scope of this paper.

cell in each market.[24]

In section 6.1, I support the informal identification arguments with Monte Carlo simulation.

5.8 Estimation

A large number of parameters in the model do not allow estimation using indirect search methods such as the method of simulated moments. To estimate the model, I use a BLP-style mapping to "invert" the observed matches and wages into the parameters (Berry et al. 1995). In each iteration, the algorithm updates the parameters based on the comparison between the predicted and actual moments.

Starting with an initial guess $(\beta^0_{txyc}, p^0_{txyc}, \sigma^0, b^0)$, the parameters are updated by the mapping

$$\beta^{h+1}_{txyc} = \beta^h_{txyc} + \eta \left[log(\mu_{txyc} \cdot w_{txyc}) - log(\hat{\mu}_{txyc}(p^h, \beta^h, \sigma^h, b^h) \cdot \hat{w}_{txyc}(p^h, \beta^h, \sigma^h, b^h))\right] \quad (23)$$

$$p^{h+1}_{txyc} = p^h_{txyc} + \eta \left[log(\mu_{txyc}) - log(\hat{\mu}_{txyc}(p^h, \beta^h, \sigma^h, b^h))\right] \quad (24)$$

$$\sigma^{h+1} = \sigma^h + \eta \left[log(WithinVar_w) - log(Within\hat{V}ar_w(p^h, \beta^h, \sigma^h, b^h))\right] \quad (25)$$

$$b^{h+1} = b^h + \eta \left[log(Var_w) - log(V\hat{a}r_w(p^h, \beta^h, \sigma^h, b^h))\right] \quad (26)$$

where $\eta > 0$ is the update rate of the parameters. The variables μ_{txyc}, w_{txyc}, $WithinVar_w$, and Var_w are the observed number of matches by a $txyc$ cell, the average wage in a cell, the between-groups wage variance, and the overall wage variance, respectively. The same variables with a "hat" are the corresponding moments predicted by the model for the parameters indicated in parentheses.[25] Finally, β^h_{txyc}, p^h_{txyc}, σ^h, and b^h are the parameters in iteration h.

To define the update equation, I use the insights about the relationships between the parameters and the moments from the previous section. Starting with the match utility parameter in equation (23), a higher utility of a specific group increases the share of matches and the average wage of that group. Therefore, both the share of matches and the average wage update this parameter. On the other hand, the meeting probability parameter posi-

[24] A key difference between the two sources of imperfect sorting is that the unobserved heterogeneity impacts only the observed sorting, but the meeting impacts both the observed and unobserved sorting. Therefore, better measures of unobserved heterogeneity might help to separately identify the two. For example, this could be done by observing the workers and firms several times. I do not explore this in the current research.

[25] To ease notation, I do not explicitly denote the dependency of the predicted moments on the idiosyncratic shocks ζ, which are fixed within the estimation. See Appendix D for additional details on the estimation.

tively impacts the share of matches but does not have a clear relationship with the average wage within a cell. Hence, it is updated only by the share of matches (equation (24)).

Two additional parameters that need to be estimated are the idiosyncratic utility parameter σ and the utility constant b, which are updated by the within-group wage variance $WithinWageVar$ and overall variance $WageVar$ (equations (25)-(26)). I add the utility constant explicitly to the estimation process, and normalize the mean of β_{txyc} (weighted by μ_{txyc}) to zero. The reason is that a naive updating of the utility parameters does not take into account the impact it has on the overall wage variation, which, in turn, could wrongly impact the estimation of σ. Updating the utility parameter location such that the total wage variance fits the actual wage variance and updating σ by the within-group wage variance directs the updating of both the utility and σ in the right direction.

6 Model results

I estimate the model 100 times with different values of the shocks ζ. In the next two sections, I present the average results (and their standard errors) across the model's 100 estimations.

6.1 Model's fit, Model's precision, and Monte Carlo simulation

Panel A of Table 6 reports measures of the model's fit to the data. The average difference (in absolute values) between the model predictions and the data is 1.3 and 0.8 log points for the matches share and average wage by a cell. The predicted wage variance and within-group wage variance are also very close to their true values, with a deviation of 0.08 and 0.07 log points. Finally, the correlation between the predicted and observed moments is almost perfect, with 1.000 for the share of matches and 0.008 for the average wage. Overall, Panel A of Table 6 shows that the model fits the data very well, which means that the BLP-style mapping successfully inverts the information on the moments into the parameters.[26]

The precision of the estimates is also high. Panel B of table 6 compares the model estimates between the 100 estimations of the model. The first row reports the average correlation in the utility and meeting parameters across any possible pair within the 100 estimations. The average correlation is 0.980 for the utility parameter and 0.988 for the meetings parameter. To check the precision of the unobserved heterogeneity, σ, and utility-

[26] This result does not say in any way that the model performs well compared to other models. A large number of parameters, which equals the number of moments, ensures that the model can fit almost any data. This check aims to show that the algorithm successfully inverts the data, although I do not have formal theoretical results to guarantee it.

scale, b, I calculate the standard deviations of their estimates across the 100 simulations. The standard deviations of $log(\sigma)$ and b are 0.007 and 0.011, which are small compared to their estimates (-1.069 and 9.174, receptively).

Finally, I investigate the identification of the model by Monte Carlo simulation. I generate data using the model, using the average parameter values described above. Pretending that the data generated by the model is the true data, I estimate the model's parameters 100 times again with different values of the shocks ζ and compare the estimates to the "true" parameters (the average over the 100 original estimates). The average correlation between each set of Monte Carlo estimates and the "true" parameters is 0.972 and 0.985 for the utility and meeting parameters, respectively (Table 6, Panel B, third row). The average estimated unobserved heterogeneity and utility-scale are -1.076 and 9.186, which are also very close to the "true" parameters, -1.069 and 9.174, respectively (Table 6, Panel B, fourth row). Overall, the results of the Monte Carlo simulation suggests that the proposed estimation procedure can identify the true parameters of the model.

6.2 Impact of parental connections

To summarize the model estimates, I run a regression on the model parameters with the workers', firms', and connections' characteristics as explanatory variables. Table 7 reports the WLS estimates of the equation

$$\theta_{txyc} = b + \delta_c + \gamma_1 Arab_x + \gamma_2 Female_x + \gamma_3 College_x + \psi_y + \epsilon_{txyc} \tag{27}$$

where each observation is weighted by the actual number of matches in the corresponding $txyc$ cell. θ_{txyc} is the parameter of interest (either match utility or meeting probability), δ_c is the connection-type effect, $Arab_x$, $Female_x$, and $College_x$ are indicators equal to one if the workers in group x are Arab, female, and college-educated, respectively, and ψ_y is the firm-type effect.

Firstly, I study the contribution of the characteristics of connections, workers, and firms to the utility parameters by estimating equation (27) with β_{txyc} as the outcome. The first column of Table 7 shows that phantom connections only slightly affect the utility parameter (1.2 log points, not statistically significantly different from 0 at the 5 percent level). Weak and strong connections increase the estimated utility by 4.1 and 15.8 log points, respectively. Taking the difference between real and phantom connections as a measure of the effect of connections, weak and strong connections increase the utility parameter by 2.8 and 14.6 log points, respectively.

The differences in firm utility from connected and not connected hiring should necessarily be interpreted as productivity differences. For example, the firm (or some workers at the firm) might benefit from hiring connected workers because of pure favoritism (or nepotism). Likewise, the firm's utility from hiring a connected worker might be higher because of a lower uncertainty about the productivity of the worker or the match. This, in turn, increases the expected time the worker will stay at the firm and therefore reduce the expected hiring, firing, or training costs. The last interpretation is consistent with the positive correlation exists in the data between connections and tenure at the first job.[27]

The coefficients of the workers' characteristics show the same sign as their sign in the wage regressions, with estimates of -1.1, -7.0, and 7.7 for Arabs, females, and college-educated workers, respectively. These coefficients represent the differences in firm-assignments and wages between new workers not explained by social connections. Other factors, such as differences in productivity, discrimination, and hours worked, might be the reason for these differences. Finally, the estimated utility is monotonically increasing with the job type, as expected.

Next, I estimate equation (27) with $log(p_{txyc})$ as an outcome. I find the effect of all types of connections on meeting probability is positive and significant (Table 7, column 2). The average meeting probability for workers and firms with phantom connections is 7.1 times higher than worker-firm pairs with no connections. The effect is stronger for firms with weak and strong connections, with an estimated 15.3 and 42.2 times higher meeting probability than unconnected pairs. Comparing phantom and real connections, weak and strong connections increase the meeting probability by factors of 2.1 and 5.9, respectively.

To further explore the model's predictions about differences in meeting probabilities for different worker groups, I run an additional regression, adding interactions between workers' characteristics and connection characteristics. Figure 6 shows the estimated meeting probabilities for each connection type by groups of ethnicity and gender. Panel A shows that the meeting probability without any connections is higher for Jews than for Arabs. However, the meeting probabilities are much higher for Arabs than for Jews for all types of connections. The difference in log points between Arabs and Jews is greater for weak and strong connections relative to phantom connections, indicating that the effect of connections is stronger for Arabs than for Jews.

[27] To separately estimate two or more of these sub-channels, a richer data is needed. For example, a direct measure of firms' profits enables isolating pure favoritism from the other channels. Likewise, dynamic information on workers and firms (accompanied by a dynamic model) can help identify the information uncertainty channel.

6.3 Sensitivity of the results to the bargaining power parameter

I estimate the benchmark model assuming a workers' bargaining power $\lambda = 0.5$. The results are not sensitive to the value of that parameter. Figure A4 plots the difference between the average estimated effects of weak connections and phantom connections on the utility and meeting parameters for different workers' bargaining power values. Starting with the match utility parameter, the estimated effects of causal weak connections (the difference between the effects of weak and phantom connections) are positive and vary between 2 and 5 log points for workers' bargaining power between 0 and 0.9, compared to 2.8 log points in the benchmark model.[28] The only exception is the unrealistic scenario when workers have perfect bargaining power. In this case, the estimated effect is very close to zero (Figure A4, Panel A).

Likewise, the estimated causal effects of weak connections on the utility parameter are not sensitive to the bargaining power parameter. The effects are between 60 and 80 log points, compared to 76 log points in the benchmark results (Figure A4, Panel B).

7 Counterfactuals

7.1 Causal connections

To get the causal effect of connections (net of the impact of confounders), I exploit the identification strategy from the first part of the paper and compare the estimated effects of real and phantom connections for each combination of workers and firms in each market. Precisely, the systematic match utility of a weak "causal" connection for workers of type x, firms of type y, and year t is

$$\beta^{causal}_{txy,weak} = \beta_{txy,none} + \beta_{txy,weak} - \beta_{txy,phantom}. \tag{28}$$

where $\beta_{txy,c}$ is the estimated systematic utility of that txy group with connections of type $c \in \{none, phantom, weak\}$. In other words, I consider the difference between the estimates of the utility with weak and phantom connections as a measure of the excess effect of connections on the utility net of confounders correlated with connections. Likewise, the

[28]Note that the value in the benchmark model is the average across 100 different estimations of the model with $\lambda = 0.5$, whereas in Figure A4 every point represent the results of a single estimation. Therefore, the value obtained in the single estimation for $\lambda = 0.5$ is not identical to the benchmark results.

meeting probability of weak "causal" connection is

$$p_{txy,weak}^{causal} = p_{txy,none} \cdot p_{txy,real} / p_{txy,phantom} \tag{29}$$

where $p_{txy,c}$ is the estimated meeting probability of that txy group with connections of type $c \in \{none, phantom, weak\}$. The analogous definitions hold for strong connections.[29]

7.2 Value of connections and meetings

In this section, I use the model to estimate the value of connections and meetings. To do so, I re-run the model with the estimated parameters and add a connection/meeting for one random pair of a worker and a firm each year. I then compare the utility of the affected workers with and without the additional connection/meeting. The utility difference measures the value of a connection or a meeting—how much the average worker will pay for one additional connection or meeting with a random firm.

I do this exercise in three ways. In first, I add a new random meeting without connections between the worker and the firm. Second, I add a new meeting and assume that the worker and firm are weakly connected. Finally, I add connections to an existing meeting. The last exercise isolates the effect of the utility channel alone.

The first column of Table 8 reports the results of this exercise with 100,000 new meetings/connections (1,000 for each of the 100 estimations of the model). For convenience, I report all results in terms of percentages of new workers' average monthly wage. The average value of one additional meeting without the utility effect is 2.2 percent of new workers' average monthly wage. Adding connections to a random existing meeting, the monthly wage increases by 1.5 percent. By combining the two effects, assuming that the new meeting is with a causal weak connected firm, the effect increases to 3.7 percent.

The model also allows decomposition of the effect into situations in which workers go to work at the firm with the new meeting/connection (with a higher wage compared to the benchmark case) and situations in which the identity of the matched firms do not change but the workers' wage increases due to the better choice set they have.

Adding a new meeting with a firm without the utility effect, in 4.0 percent of the cases the worker is matched with that new firm. The average gains are 41.4 percent of the average wage. In 6.4 percent of the cases, the new meeting does not lead to a new job but increases the salary due to that worker's better choice set. The average gains, in that case, are 7.9

[29] Because the accuracy of the estimates of cells with no or a small number of matches is low, I censor the extreme values of the parameters in the calculation. See Appendix D for the exact definitions.

percent of the average wage (Table 8, row 1).

If we add the utility effect of causal weak connections to existing meetings, in 4.0 percent of the cases, the worker changes her job to a new connected job. The average gains are 20.3 percent of the average wage, so the expected gains are 0.8 percent. In 10.1 percent of the cases, the wage changes without a job change, with expected gains of 6.4 percent of the average wage (Table 8, row 2).

Finally, if we assume that the new meeting is accompanied by the utility of a causal weak connection, the probability that the workers end up working at the new firm is 5.5 percent. In this case, the average gains are 57 percent of the average wage, and the contribution of this event to the total gains is 3.1 percent of the average wage. In 5.8 percent of the cases, the wage changes without a job change. These events yield average gains of 9 percent of the average wage (Table 8, row 3).

The decomposition of the contribution of events with and without job changes shows that about 84 percent of the value of connections comes from a direct effect of the new meeting/connection that leads to a better job with a better salary. However, an indirect effect, namely the impact of the new meeting/connection on the salary through a better choice set of the worker, makes a non-negligible contribution to the overall value.

Not all meetings/connections are equal. Figure 7 shows the expected effect by the job type of the new meeting/connection. The results indicate that having a new meeting with a high-ranked firm (i.e., a firm in the upper quintile of AKM firm premium) is much more valuable than a meeting with a lower-ranked firm. This result is true in all scenarios (a new meeting without the utility effect, an existing meeting with the utility effect, and a new meeting with the utility effect).

7.3 Between-group pay gaps

How much of the pay gap between different groups in Israel is due to differences in social capital? I use the structural model to answer that question in two ways. Firstly, I check the predicted inequality if the different groups, Arabs and Jews or males and females, would have similar quantities and qualities of connections. Secondly, I check the predicted pay gaps given a policy that prohibits using different types of social connections.

I perform the first exercise by adding random connections to workers such that the number of weak and strong connections per worker with each firm type is equal between the groups. For example, for the ethnicity characteristic, I compare the number of meetings per worker for Jews and Arabs in the same year, with the same gender and education characteristics, and with the same type of firm. Then, I add random connections of that type to the group with fewer connections until the number of connections per worker is equal.

To see the importance of my identification strategy—evaluating the effects of connections by comparing real and phantom connections—I check the model's predictions with and without that strategy. Without the identification strategy, the meeting and utility parameters of a new connection are simply the parameters of real connections (either weak or strong) of that $txyc$ cell. With the identification strategy, I assume that the new connections have only the excess impact of real connections relative to phantom connections, as defined above (equation 28 and 29 and the analogous definition for strong connections).

Starting with the ethnic pay gap, the first row of Table 9 shows the results when the share of connections with all firms is equal for Arabs and Jews. The benchmark gap in monthly wage between Arabs and Jews is 502 Shekels, or 8.4 percent of the average monthly wage. Without the identification strategy, we wrongly attribute the impact of confounders, which are correlated with connections, to the effect of connections themselves, and therefore incorrectly concluding that parental connections explain an unrealistically large fraction of the ethnic wage gap.

The gap estimates are much closer to the benchmark gap when correctly using the identification strategy. The estimated reduction in the ethnicity pay gap is now 5, 2, and 12 percent, given the meeting effect, utility effect, and both effects, respectively. The large difference between the counterfactual results with and without the identification strategy indicates the importance of using identification variation in structural estimation and interference. Without the identification strategy, we wrongly attribute the impact of confounders, which are correlated with connections, to the effect of connections themselves; therefore, obtaining that parental connections explain a non-realistic large fraction of the ethnic wage gap (Table 9, Panel A, first row).

In contrast to the ethnic pay gap, equalizing males' and females' parental connections has no significant effect on the gender wage gap. Without the identification strategy, the counterfactual gender pay gap is increasing by 2.3 percent. However, using the identification strategy, the gap increases by 0.1 percent, and the change is not significantly different from zero (Table 9, Panel A, second row).

Next, I check the counterfactual pay gaps under the assumption that hiring a worker with real connections is forbidden. I check the effect of this policy for weak connections only, for strong connections only, and for strong and weak connections together. Panel B of Table 9 shows that prohibiting the hiring of workers with connections, as some anti-nepotism rules do, increases the predicted ethnic pay gap by 8 percent if only weak connections are prohibited, 44 percent if only strong connections are prohibited, and 56 percent if both weak and strong connections are prohibited. The gender pay gap declines by 4, 20, and 25 percent respectively in these different scenarios.

The difference between the results of the two different scenarios can be explained by considering both the differences in the quality of connections and in the "return" to connections of the different groups. For example, the model predicts that equalizing the connections between Arabs and Jews reduces the ethnic pay gap, but prohibiting connections increases it. The explanation for this comes from two opposing forces. On the one hand, Arabs have worse connections in the labor market compared to Jews (Table 1 and Figure 1). On the other hand, Arabs rely heavily on connections compared to Jews, as can be seen from the higher effect of connections both in the reduced-form and structural estimates (Figures 4 and 6 4). Therefore, equalizing Arabs and Jews' connections provides them better connections, which reduces the pay gap. However, prohibiting the use of connections increases the gap as it hurts Arabs more than Jews and, therefore, increases the gap. The results of the gender gap are different. As there is no big difference between the parental connections of males and females, equalizing the connections does not impact the gender gap. However, because the return to connections is higher for males than females, prohibiting connections hurts males more than females and reduces the gap.

8 Conclusion

In this paper, I study the role of parental social networks in shaping the distribution of job assignments and the wages of new workers. To do so, I leverage the timing of between-job moves of potential contacts relative to the labor-market entry year for exogenous variation of the social networks. In the first part of the paper, I use regression analysis to estimate the effect of strong and weak parental connections on job assignments. Then, I build and estimate a matching model with search frictions where heterogeneous workers and firms choose their best match given their choice set and the set of wages that clear the market. I allow social connections to impact both the available choice sets and the match values.

In the reduced-form part, I find that workers are 3-4 times more likely to find employment in firms where a past coworker of the parent currently works than in otherwise similar firms. I show that the effect is more potent if the potential connections are formed in smaller firms or more recently. I also find a positive correlation between the wage of new workers and parental connections.

A structural estimation of the model shows that parental connections increase the meeting probability and the potential match value. Exploiting the same identification strategy, I find that a weakly connected worker-firm pair is twice as likely to meet than a phantom-connected pair. Likewise, the match value is higher by 2.8 percent for weakly versus phantom connected pairs. Using the model estimates, I find that workers are willing to pay, on average, 3.7 percent of the average wage to get one additional meeting with a connected firm. I also find that differences in parental network quality explains a large proportion of Israel's ethnic pay gap. Equalizing the quantity and quality of Arabs' and Jews' connections, the ethnic pay gap decreases by 12 percent. However, because Arabs rely more than Jews on connected hiring, prohibiting the hiring of connected workers increases the gap by 56 percent.

I make three methodological contributions. Firstly, I separately estimate the two major impacts described in the literature of social connections on labor market outcomes: search and match quality. Secondly, I contribute to the two-sided matching literature by introducing search frictions into this type of model. I exploit the assignment problem's sparsity implied by this assumption, together with recent developments in assignment problem algorithms, to simulate the model. Thus, I can perform a simulation-based estimation of the model, which allows for potentially rich and flexible utility functions. To the best of my knowledge, this is the first simulation-based estimation of a matching model with large scale data. Thirdly, I suggest a novel "BLP-style" estimation procedure to estimate two sets of unobserved choice characteristics with two sets of data points. In each iteration, the parameters are updated one by one to the direction that best fits the data. This direct updating procedure enables estimation of models with many parameters, even when simulation of each model's iteration is expensive. Taken together, the model proposed in this paper can serve as a workhorse for studying various questions regarding the labor market.

My empirical results have nuanced consequences for policymakers. Policies to reduce the inequality implied by differential parental networks include, for example, subsidies for internships in good firms for graduates from lower socioeconomic backgrounds, or "Rooney Rule" policies requiring interview of these candidates for open positions (Solow et al. 2011). The results of the model also shed light on the expected outcomes of different policies. For instance, a long-term internship is likely to impact not only the "search frictions" (e.g., the probability for a job-interview at the firm), but also the "match value", through better information on the workers and match quality. On the other hand, "Rooney Rule"-type policies are likely to impact only the "search frictions" and therefore have a more moderate effect on inequality. Finally, the model suggests that policies that entirely prohibit the use of connections might have the opposite effect on inequality, as workers from lower socioeconomic backgrounds rely more on social links in the labor market.

The framework employed here can be readily ported to other datasets and problems, and there is ample room for future research. Firstly, like most of the matching literature, the model is static. Estimating a dynamic version of the model will enable study of how connections matter over the life cycle and explicit modeling of the impact of referrals on the firm's uncertainty about worker quality. Additionally, observing the same workers several times allows estimation of workers' and firms' fixed effects, which cannot be separately identified in a static model. Secondly, having information on other labor market outcomes could allow the estimation of additional unobserved parameters, such as the workers' non-wage match utility and differential workers' bargaining power. Such data include direct information on firms' production or the meeting/interview process. Further unpacking the black box of the matching between workers and firms is an essential step in crafting policies to help reduce inequity and increase efficiency in the labor market.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, "High wage workers and high wage firms," *Econometrica*, 1999, *67* (2), 251–333.

Agarwal, Nikhil, "An empirical model of the medical match," *American Economic Review*, 2015, *105* (7), 1939–78.

Ahyja, Ravindra K., James B. Orlin, and Thomas L. Magnanti, *Network flows: theory, algorithms, and applications*, Prentice-Hall, 1993.

Arellano-Bover, Jaime and Shmuel San, "The Role of Firms in the Assimilation of Immigrants," *Unpublished*, 2020.

Athey, Susan, Christopher Avery, and Peter Zemsky, "Mentoring and diversity," *American Economic Review*, 2000, *90* (4), 765–786.

Azoulay, Pierre, Joshua S. Graff Zivin, and Jialan Wang, "Superstar extinction," *The Quarterly Journal of Economics*, 2010, *125* (2), 549–589.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, "Social connections and incentives in the workplace: Evidence from personnel data," *Econometrica*, 2009, *77* (4), 1047–1094.

Banfi, Stefano and Benjamin Villena-Roldan, "Do high-wage jobs attract more applicants? Directed search evidence from the online labor market," *Journal of Labor Economics*, 2019, *37* (3), 715–746. Publisher: The University of Chicago Press Chicago, IL.

Beaman, Lori and Jeremy Magruder, "Who gets the job referral? Evidence from a social networks experiment," *American Economic Review*, 2012, *102* (7), 3574–93.

Bernard, Florian, Nikos Vlassis, Peter Gemmar, Andreas Husch, Johan Thunberg, Jorge Goncalves, and Frank Hertel, "Fast correspondences for statistical shape models of brain structures," in "Medical Imaging 2016: Image Processing," Vol. 9784 International Society for Optics and Photonics 2016, p. 97840R.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, 1995, pp. 841–890.

Bertsekas, Dimitri P., *Network optimization: continuous and discrete models*, Athena Scientific Belmont, MA, 1998.

Boca, Daniela Del and Christopher J. Flinn, "Household behavior and the marriage market," *Journal of Economic Theory*, 2014, *150*, 515–550.

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

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa, "A distributional framework for matched employer employee data," *Econometrica*, 2019, *87* (3), 699–739. ISBN: 0012-9682 Publisher: Wiley Online Library.

Bonnet, Odran, Alfred Galichon, Keith O'Hara, and Matthew Shum, "Yogurts Choose Consumers? Estimation of Random Utility Models via Two-Sided Matching," *Unpublished*, 2018.

Caldwell, Sydnee and Nikolaj Harmon, "Outside Options, Bargaining, and Wages: Evidence from Coworker Networks," *Unpublished*, 2018, p. 107.

_ and Oren Danieli, "Outside options in the labor market," *Unpublished*, 2018.

Calvo-Armengol, Antoni and Matthew O. Jackson, "The effects of social networks on employment and inequality," *American economic review*, 2004, *94* (3), 426–454.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women," *The Quarterly Journal of Economics*, 2015, *131* (2), 633–686.

_ , _ , Jörg Heining, and Patrick Kline, "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

_ , Jörg Heining, and Patrick Kline, "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly journal of economics*, 2013, *128* (3), 967–1015.

Chiappori, Pierre-André and Bernard Salanié, "The econometrics of matching models," *Journal of Economic Literature*, 2016, *54* (3), 832–61.

Choo, Eugene and Aloysius Siow, "Who marries whom and why," *Journal of political Economy*, 2006, *114* (1), 175–201.

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

Cohen, Asher and Bernard Susser, "Jews and Others: Non-Jewish Jews in Israel," *Israel Affairs*, 2009, *15* (1), 52–65.

Corak, Miles and Patrizio Piraino, "The intergenerational transmission of employers," *Journal of Labor Economics*, 2011, *29* (1), 37–68.

Demange, Gabrielle and David Gale, "The strategy structure of two-sided matching markets," *Econometrica: Journal of the Econometric Society*, 1985, pp. 873–888. Publisher: JSTOR.

Dickinson, David L., David Masclet, and Emmanuel Peterle, "Discrimination as favoritism: The private benefits and social costs of in-group favoritism in an experimental labor market," *European Economic Review*, 2018, *104*, 220–236.

Dupuy, Arnaud and Alfred Galichon, "Personality traits and the marriage market," *Journal of Political Economy*, 2014, *122* (6), 1271–1319.

_ **and** _ , "A Note on the Estimation of Job Amenities and Labor Productivity," *Unpublished*, 2015.

Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker, "Referral-based job search networks," *The Review of Economic Studies*, 2016, *83* (2), 514–546.

Eliason, Marcus, Lena Hensvik, Francis Kramarz, and Oskar Nordstrom Skans, "Social Connections and the Sorting of Workers to Firms," *Unpublished*, 2019.

Fontaine, Francois, "Why are similar workers paid differently? The role of social networks," *Journal of Economic Dynamics and Control*, 2008, *32* (12), 3960–3977. Publisher: Elsevier.

Galichon, Alfred and Bernard Salanié, "Cupid's invisible hand: Social surplus and identification in matching models," *Unpublished*, 2015.

Galor, Oded and Daniel Tsiddon, "Technological progress, mobility, and economic growth," *The American Economic Review*, 1997, pp. 363–382.

Gerard, François, Lorenzo Lagos, Edson Severnini, and David Card, "Assortative matching or exclusionary hiring? the impact of firm policies on racial wage differences in brazil," *Unpublished*, 2018. ISBN: 0898-2937.

Granovetter, Mark, *Getting a job: A study of contacts and careers*, University of Chicago press, 1973.

Gretsky, Neil E., Joseph M. Ostroy, and William R. Zame, "Perfect competition in the continuous assignment model," *Journal of Economic Theory*, 1999, *88* (1), 60–118.

Jaffe, Sonia and Simon Weber, "The effect of meeting rates on matching outcomes," *Economic Theory*, 2019, *67* (2), 363–378. Publisher: Springer.

Jager, Simon, "How substitutable are workers? evidence from worker deaths," *Unpublished*, 2016.

Kramarz, Francis and David Thesmar, "Social networks in the boardroom," *Journal of the European Economic Association*, 2013, *11* (4), 780–807.

_ **and Oskar Nordström Skans**, "When strong ties are strong: Networks and youth labour market entry," *Review of Economic Studies*, 2014, *81* (3), 1164–1200.

Montgomery, James D., "Social networks and labor-market outcomes: Toward an economic analysis," *The American economic review*, 1991, *81* (5), 1408–1418.

Mortensen, Dale, *Wage dispersion: why are similar workers paid differently?*, MIT press, 2003.

Mortensen, Dale T. and Christopher A. Pissarides, "Job creation and job destruction in the theory of unemployment," *The review of economic studies*, 1994, *61* (3), 397–415. Publisher: Wiley-Blackwell.

Noy, Chaim and Erik Cohen, *Israeli backpackers: From tourism to rite of passage*, SUNY Press, 2005.

of Statistics, Central Bureau, *Statistical abstract of Israel*, Jerusalem, Israel, 2019.

Plug, Erik, Bas van der Klaauw, and Lennart Ziegler, "Do Parental Networks Pay Off? Linking Children's Labor-Market Outcomes to Their Parents' Friends," *The Scandinavian Journal of Economics*, 2018, *120* (1), 268–295.

Postel–Vinay, Fabien and Jean-Marc Robin, "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica*, 2002, *70* (6), 2295–2350.

Salanié, Bernard, "Identification in Separable Matching with Observed Transfers," *Unpublished*, 2015.

Schmutte, Ian M., "Job referral networks and the determination of earnings in local labor markets," *Journal of Labor Economics*, 2015, *33* (1), 1–32.

Shapley, Lloyd S. and Martin Shubik, "The assignment game I: The core," *International Journal of game theory*, 1971, *1* (1), 111–130.

Solow, Benjamin L., John L. Solow, and Todd B. Walker, "Moving on up: The Rooney rule and minority hiring in the NFL," *Labour Economics*, 2011, *18* (3), 332–337. Publisher: Elsevier.

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter, "Firming up inequality," *The Quarterly journal of economics*, 2019, *134* (1), 1–50.

Topa, Giorgio, "Labor markets and referrals," in "Handbook of social economics," Vol. 1, Elsevier, 2011, pp. 1193–1221.

Tables and Figures

Table 1: Summary statistics—new workers

	All	Ethnicity		Gender	
		Jews	Arabs	Males	Females
N.	220,806	157,023	63,783	126,233	94,573
Arabs	0.29	0.00	1.00	0.37	0.19
Females	0.43	0.49	0.28	0.00	1.00
College	0.23	0.25	0.16	0.15	0.33
First job					
Age	24.00	24.22	23.48	23.82	24.25
Salary	5,839	6,053	5,312	6,223	5,325
Tenure	2.01	1.97	2.10	2.04	1.98
Firm rank	0.60	0.64	0.52	0.60	0.61
Connections					
Weak	0.03	0.02	0.04	0.03	0.02
Strong	0.11	0.09	0.17	0.13	0.08
Age 30					
Salary	8,939	9,373	7,317	9,806	7,832
Firm rank	0.68	0.70	0.58	0.67	0.68
Connections					
Av. firm rank					
Weak	0.64	0.66	0.58	0.63	0.65
Strong	0.61	0.64	0.54	0.60	0.62
N. firms					
Weak	43.66	50.40	26.78	41.71	46.26
Strong	24.41	27.25	17.39	23.70	25.34

Notes: This table reports summary statistics for the sample of new workers. The first column reports the average value of the variables for the entire sample, and the other columns report for sub-samples separated according to ethnicity and gender. Firm rank is the rank of the firm-specific pay premium estimated using an AKM model Abowd et al. (1999). "Connections" indicates whether the worker has weak or strong connections at the first job. Av. firm rank of connections is the average firm rank of firms with which the worker has weak and strong connections. N. firms is the number of such firms.

Table 2: Effects of parental connections on firm assignment

	All	Jews	Arabs	Males	Females
	(1)	(2)	(3)	(4)	(5)
Phantom connections	0.010	0.006	0.030	0.011	0.008
	[0.009,0.011]	[0.005,0.007]	[0.025,0.032]	[0.010,0.013]	[0.006,0.010]
Weak connections	0.050	0.031	0.143	0.067	0.031
	[0.047,0.054]	[0.028,0.034]	[0.131,0.156]	[0.061,0.071]	[0.027,0.036]
Strong connections	0.487	0.366	0.917	0.617	0.338
	[0.472,0.501]	[0.351,0.384]	[0.878,0.956]	[0.593,0.647]	[0.320,0.354]
R0 (no connections)	0.005	0.005	0.006	0.005	0.006
	[0.005,0.005]	[0.005,0.005]	[0.006,0.006]	[0.005,0.005]	[0.005,0.006]
Ratio weak-phantom	3.666	3.259	4.177	4.409	2.731
	[3.316,4.081]	[2.841,3.681]	[3.651,4.803]	[3.912,4.959]	[2.262,3.303]
Ratio strong-phantom	32.52	33.99	25.91	38.37	25.37
	[30.02,35.53]	[30.65,37.8]	[23.52,30.03]	[34.83,43.67]	[22.41,29.39]
Observations	21,166,443	16,837,526	4,328,917	15,319,313	5,847,130
N firms	149,729	144,186	117,746	145,939	134,555
N groups	2,959	1,658	1,301	1,548	1,411
N workers	220,684	157,009	63,675	170,872	49,812
N connections	40,827,833	33,261,814	7,566,019	31,664,340	9,163,493

Notes: This table reports the probability of working in a firm with different types of connections, relative to working in a non-connected firm. The coefficients are the mean coefficients of phantom, weak, and strong connections across 100 estimations of equation (4) using a 20 percent random sample of workers each time. I construct the bounds of the 95 percent confidence intervals using the 2.5 and 97.5 percentiles of the coefficients' distribution. R0 is the average probability of working in a non-connected firm. "Ratio weak-phantom" is the estimated odds ratio between working at a weakly-connected firm and working in a phantom-connected firm. "Ratio strong-phantom" is defined similarly. The first column reports the results for the entire sample, while the other columns report the results for a different sub-group of the new workers each time.

Table 3: Effects of parental connections on firm assignment: death and retirement of contacts

	Employment		
	(1)	(2)	(3)
Special connections:	Death	Retirement	Death or retirement
Phantom (D/R)	0.031	0.010	0.017
	[0.004,0.068]	[-0.008,0.032]	[0.001,0.034]
Phantom (Other)	0.010	0.010	0.010
	[0.009,0.011]	[0.009,0.011]	[0.009,0.011]
Weak (D/R)	0.065	0.032	0.041
	[0.010,0.126]	[0.003,0.066]	[0.017,0.071]
Weak (Other)	0.050	0.051	0.051
	[0.047,0.054]	[0.047,0.055]	[0.047,0.054]
Strong	0.487	0.487	0.487
	[0.472,0.501]	[0.472,0.501]	[0.472,0.501]
R0 (no connections)	0.005	0.005	0.005
	[0.005,0.005]	[0.005,0.005]	[0.005,0.005]
Ratio weak-phantom (D/R)	2.567	3.913	2.773
	[0.386,7.746]	[0.582,19.460]	[0.748,6.533]
Ratio weak-phantom (Other)	3.679	3.680	3.691
	[3.335,4.101]	[3.339,4.099]	[3.349,4.122]
N connections: phantom (D/R)	85,532	138,194	222,461
N connections: weak (D/R)	37,402	102,499	138,974

Notes: This table reports the probability of working in a firm with different types of connections, relative to working in a non-connected firm. I divide phantom and weak connections into "D/R" connections ("death", "retirement" or both, depending on the column) and "Other" connections. "Death" connections are connections in which the contact died no more than one year after the last year she worked at the firm. "Retirement" connections are connections in which the last year the contact worked at the firm was at the mandatory retirement age (62 for females and 67 for males). In the third column, I use either death or retirement connections. The coefficients are the mean coefficients across 100 estimations of equation (4) with separate coefficients for "special" and "other" phantom and weak connections and using a 20 percent random sample of workers each time. I construct the bounds of the 95 percent confidence intervals using the 2.5 and 97.5 percentiles of the coefficients' distributions. R0 is the average probability of working in a non-connected firm. "Ratio weak-phantom: D/R" is the estimated odds ratio between working at a special weakly-connected firm and working in a special phantom-connected firm. "Ratio weak-phantom: Other" is defined similarly.

Table 4: Effect of weak parental connections on firm assignment, placebo test

	All	Jews	Arabs	Males	Females
	(1)	(2)	(3)	(4)	(5)
Phantom connections	0.000	0.000	0.000	0.000	0.000
	[-0.001,0.001]	[-0.001,0.001]	[-0.002,0.003]	[-0.001,0.001]	[-0.001,0.001]
Weak connections	0.000	0.000	0.000	0.000	0.000
	[-0.002,0.002]	[-0.002,0.002]	[-0.006,0.006]	[-0.002,0.003]	[-0.003,0.003]
Strong connections	0.000	0.000	0.001	0.000	0.000
	[-0.006,0.007]	[-0.005,0.005]	[-0.021,0.021]	[-0.006,0.008]	[-0.008,0.010]
R0 (no connections)	0.007	0.006	0.011	0.008	0.007
	[0.007,0.008]	[0.006,0.007]	[0.011,0.012]	[0.007,0.008]	[0.007,0.007]
Ratio weak-phantom	1.010	1.000	1.053	1.011	1.017
	[0.755,1.384]	[0.727,1.330]	[0.397,1.645]	[0.660,1.334]	[0.631,1.524]
Ratio strong-phantom	1.047	1.029	1.107	1.065	1.036
	[0.206,2.019]	[0.189,1.805]	[-0.938,3.233]	[0.154,1.981]	[-0.162,2.471]
Observations	21,166,443	16,837,526	4,328,917	15,319,313	5,847,130
N firms	149,729	144,186	117,746	145,939	134,555
N groups	2,959	1,658	1,301	1,548	1,411
N workers	220,684	157,009	63,675	170,872	49,812
N connections	40,827,833	33,261,814	7,566,019	31,664,340	9,163,493

Notes: This table shows placebo test results, assigning the worker's connections to a random worker in her group. The table reports the probability of working in a firm with different types of connections, relative to working in a non-connected firm, based on the new (randomized) data. The coefficients are the mean coefficients of phantom, weak, and strong connections across 100 estimations of equation (4) using a 20 percent random sample of workers each time. I construct the bounds of the 95 percent confidence intervals using the 2.5 and 97.5 percentiles of the coefficients' distributions. R0 is the average probability of working in a non-connected firm. "Ratio weak-phantom" is the estimated odds ratio between working at a weakly-connected firm and working in a phantom-connected firm. "Ratio strong-phantom" is defined similarly. The first column reports the results for the entire sample, while the other columns report the results for a different sub-group of the new workers each time.

Table 5: Correlation between parental connections at first job and salary and tenure

	Log salary	Job tenure
	(1)	(2)
Phantom connections	0.012	0.098
	(0.004)	(0.022)
Weak connections	0.026	0.187
	(0.004)	(0.025)
Strong connections	0.083	0.441
	(0.003)	(0.020)
Group FE	Yes	Yes
Firm FE	Yes	Yes
Observations	220,806	220,806
N firms	54,321	54,321
R^2 (full model)	0.624	0.414
R^2 (projected model)	0.006	0.007

Notes: This table reports the correlation between parental connections of different types and the salary and tenure at the first job. The outcome variable in the first column is (log) monthly salary in the first year of the first job. The outcome variable in the second column is the number of sequential years the worker worked at the first job (truncated at 2015). The two specifications include group and firm fixed effects. Groups are constructed using all combinations of the workers' observable characteristics (ethnicity, education, gender, year of first job, age, and district of residence). Robust standard errors clustered by group and firm are reported in parentheses.

Table 6: Model's fit, Model's precision, and Monte Carlo simulation

A. Model's fit

	Matches (μ_{txyc})	Av. wage (w_{txyc})	Overall wage variance	Within-group wage variance
	(1)	(2)	(3)	(4)
Abs. deviation	0.013	0.008	0.0008	0.0007
	(0.0006)	(0.0006)	(0.0006)	(0.0005)
Correlation	1.000	0.998		
	(0.00002)	(0.0002)		

B. Model's precision and Monte Carlo simulation

	Utility (β_{txyc})	Meetings (p_{txyc})	Unobserved heterogeneity ($log(\sigma)$)	Utility scale (b)
	(1)	(2)	(3)	(4)
Estimates				
Correlation	0.980	0.988		
	(0.001)	(0.0006)		
Value			-1.069	9.174
			(0.007)	(0.011)
Monte Carlo				
Correlation	0.972	0.985		
	(0.003)	(0.0006)		
Value			-1.076	9.186
			(0.006)	(0.009)

Notes: This table reports measures of the model's fit to the data (Panel A), the model's precision, and the results of Monte Carlo simulation (Panel B). The first row reports the average difference between the predicted and true moments on a logarithmic scale (averaged over all $txyc$ cells with weights equal to the observed matches in each cell μ_{txyc} in the first two columns). The second row of Panel A shows the correlation between the true and predicted moments (with the same weights). Each statistic in Panel A is calculated separately for each of the 100 estimations of the model, and the table reports the averages across the 100 estimations (and their standard errors in parentheses). The first row of Panel B reports the average correlation in the utility and meeting parameters across any possible pair within the 100 estimations (and their standard errors in parentheses). The second row reports the average values (and standard errors) of the unobserved heterogeneity $/sigma$, and utility-scale b parameters across the 100 simulations. The last two rows report the results of Monte Carlo simulation, where I use the average parameter values as the "true parameters" to generate data and estimate the model 100 times again with different idiosyncratic shocks. The third row reports the average correlation in the utility and meeting parameters between the new estimates and the "true parameters". The final row shows the average value of the other two parameters. Standard errors across the 100 Monte Carlo estimations are reported in parentheses.

Table 7: Projection of the model estimates on workers', firms', and connections' characteristics

	Firm's utility (β_{txyc})	Meeting probability ($Log(p_{txyc})$)
	(1)	(2)
Constant	8.809	-6.900
	(0.011)	(0.015)
Phantom connections	0.012	1.964
	(0.007)	(0.039)
Weak connections	0.041	2.728
	(0.008)	(0.038)
Strong connections	0.158	3.742
	(0.004)	(0.019)
Arab	-0.011	0.051
	(0.002)	(0.010)
Female	-0.070	-0.009
	(0.002)	(0.010)
College	0.077	-0.066
	(0.002)	(0.011)
Job type: 2	0.120	-0.067
	(0.005)	(0.012)
Job type: 3	0.268	-0.028
	(0.005)	(0.012)
Job type: 4	0.459	-0.002
	(0.006)	(0.013)
Job type: 5	0.967	-0.093
	(0.007)	(0.021)
Weak - phantom	0.028	0.764
	(0.010)	(0.054)
Strong - phantom	0.146	1.779
	(0.008)	(0.042)
R^2	0.907	0.831
	(0.003)	(0.005)

Notes: This table reports the results of regressing the utility and meeting parameters obtained from the model on worker, firm, and connection characteristics. I estimate the regression using weighted least squares, with weights equal to the actual number of matches of the *txyc* cell. "Weak (Strong) - phantom" is the difference between the coefficients of weak (strong) and phantom connections. Each regression is calculated separately for each of the 100 estimations of the model, and the table reports the averages across the 100 estimations (and their standard errors in parentheses).

Table 8: Value of meetings and connections

	Total expected gains	Salary change with a job change			Salary change without a job change		
		Probability	Gains	Expected gains	Probability	Gains	Expected gains
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
New meeting, without utility effect	2.2	0.040	41.4	1.7	0.064	7.9	0.5
	(0.417)	(0.007)	(6.543)	(0.394)	(0.008)	(1.809)	(0.135)
Existing meeting, with utility effect	1.5	0.040	20.3	0.8	0.101	6.4	0.7
	(0.467)	(0.007)	(8.151)	(0.373)	(0.010)	(2.974)	(0.311)
New meeting, with utility effect	3.7	0.055	57.0	3.1	0.066	9.0	0.6
	(0.819)	(0.009)	(9.323)	(0.778)	(0.008)	(2.248)	(0.153)

Notes: This table shows the impact of a new meeting or connection on the average worker's expected value. Each row reports the average change in the salary of workers in 5000 simulations of 3 different scenarios: 1) adding a meeting to a random worker and firm in each market, assuming no connections between them, 2) choosing a random non-connected pair in each market and changing the systematic match utility to reflect the utility of a causal weak connection, and 3) adding a random meeting with causal weak connections. The utility of a causal weak connection is the excess utility of weak connections compared to phantom connections. The first column reports the total expected gains. In the rest of the columns, I decompose that effect into two events. In columns (2)-(4), the new meeting or connection impacts the identity of the firm the worker ends up working at (compared to the job before the change). In the last three columns, the worker stays in the same position with and without the shock, but her salary changes due to a change in the available choice set. For each event, I report the probability of this event to happen, the average gains, and the expected gains of this event (probability multiplied by gains). Each statistic is calculated separately for each of the 100 estimations of the model, and the table reports the averages across the 100 estimations (and their standard errors in parentheses).

Table 9: Counterfactual impacts of connections on between-group pay gaps

A. Equalizing number of connections per worker

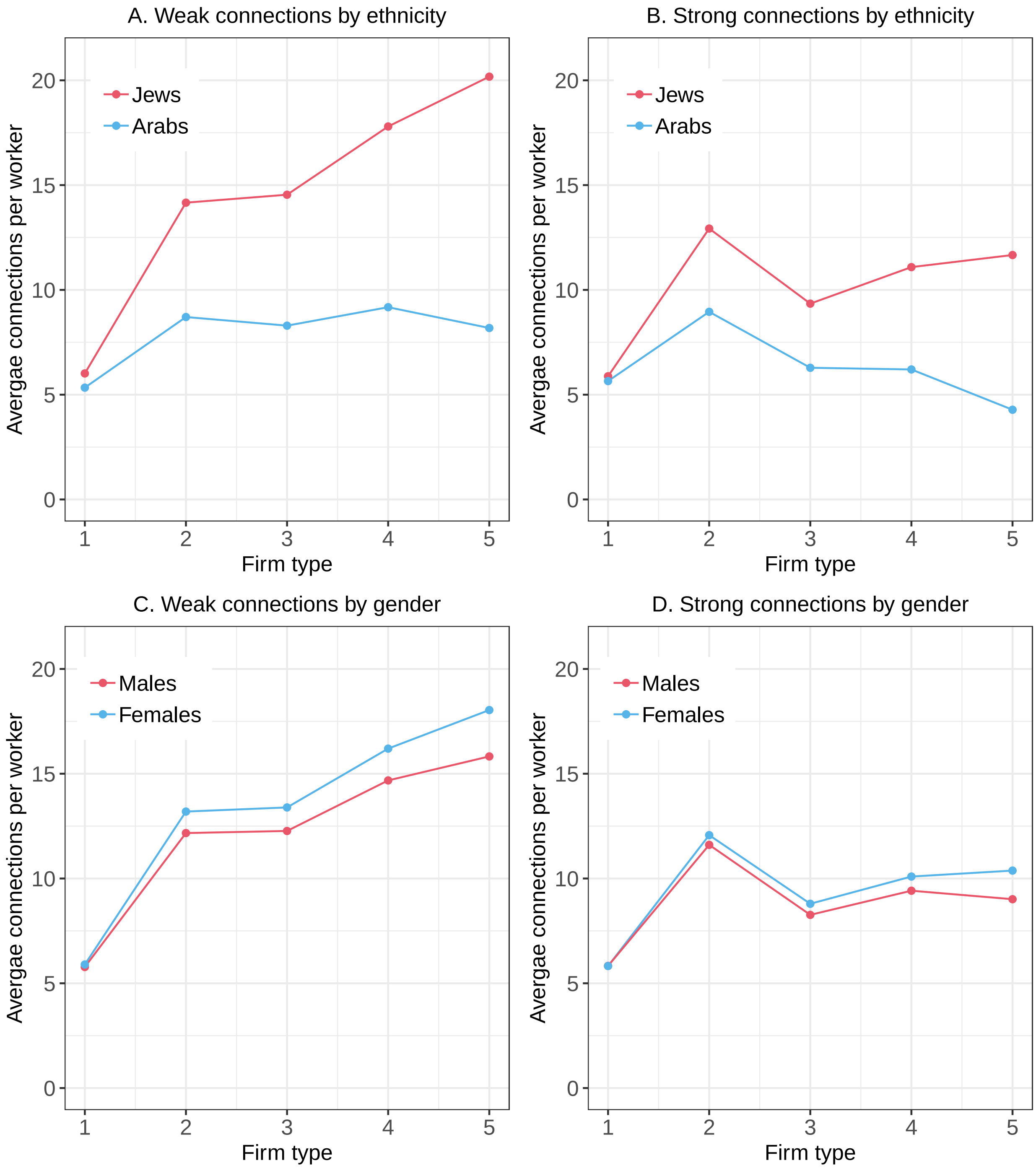
	Gap	Without identification strategy			With identification strategy		
	(% Average)	Meetings effect	Utility effect	Both effects	Meetings effect	Utility effect	Both effects
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethicity gap	-8.4	-59.5	-0.4	-67.6	-5.1	-1.1	-11.7
	(0.351)	(4.866)	(0.168)	(3.031)	(0.679)	(0.297)	(1.638)
Gender gap	-18.0	1.2	0.0	2.3	0.1	0.0	0.1
	(0.290)	(0.180)	(0.034)	(0.197)	(0.066)	(0.045)	(0.093)

B. Prohibiting hiring of connected workers

	Baseline	Weak	Strong	Weak + strong
	(% Average)			
	(1)	(2)	(3)	(4)
Ethnicity gap	-8.4	8.9	44.3	56.4
	(0.351)	(0.982)	(2.820)	(3.347)
Gender gap	-18.0	-4.0	-20.3	-25.3
	(0.290)	(0.320)	(0.780)	(0.798)

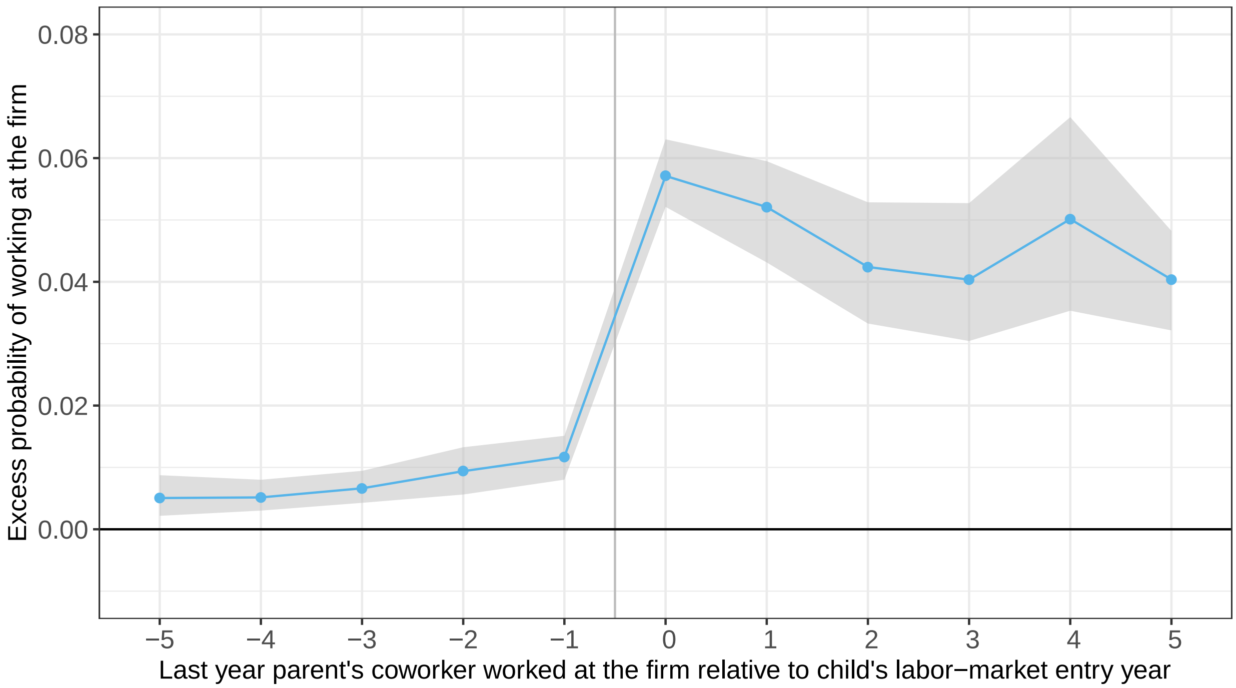
Notes: This table shows the contribution of parental connections to the ethnic and gender pay gaps in two scenarios. Panel A reports estimates from equalizing the connections between the ethnic and gender group. Specifically, in the first row, I present the ethnic pay gap predicted by the model assuming each group of Arabs and Jews (with similar gender and education characteristics) have the same number of weak and strong connections per worker with every type of firm. The second row reports the analogous results for the gender gap. Column (1) reports the benchmark pay gap as a share of the average wage. In columns (2)-(5), I estimate the counterfactual pay gaps under the assumption that new (causal) connections (either weak or strong) have the same impact on the meeting rate and the match utility as a real connection of the same type in the same *txyc* cell. In columns (6)-(8), I assume that the impact of a new (causal) connections on the meeting rate and the match's utility is the excess impact of strong or weak connections on these parameters compared to phantom connections. In columns (2) and (5), I shut down the utility effect of new connections (assuming they are similar to the utility of that *txyc* group without connections) to examine the impact of the meeting rate alone. Similarly, in columns (2) and (5), I shut down the meetings effect. In columns (4) and (7), I estimate the ethnic wage gap with both effects. Panel B reports the estimated gaps from the scenario that hiring of connected workers is prohibited. Columns (2), (3), and (4) assume hiring of workers with weak, strong, or either is banned, respectively. Each statistic is calculated separately for each of the 100 estimations of the model, and the table reports the averages across the 100 estimations (and their standard errors in parentheses).

Figure 1: Average connected firms per worker by worker characteristics, firm type, and connection type



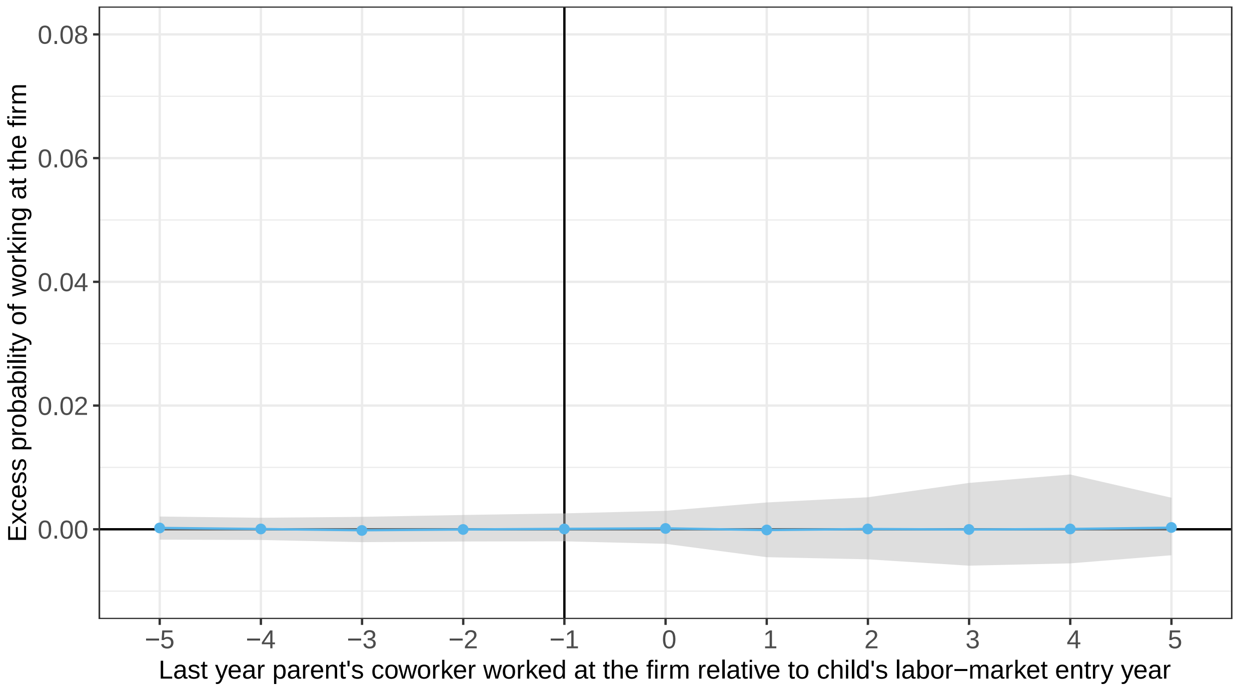
Notes: This figure shows the average number of weakly and strongly connected firms per worker by workers' ethnicity, education, and gender, and by quintiles of the AKM firm premium averaged over the years 2006-2015.

Figure 2: Event-study plot of coefficients: Effect of weak parental connections on firm assignment



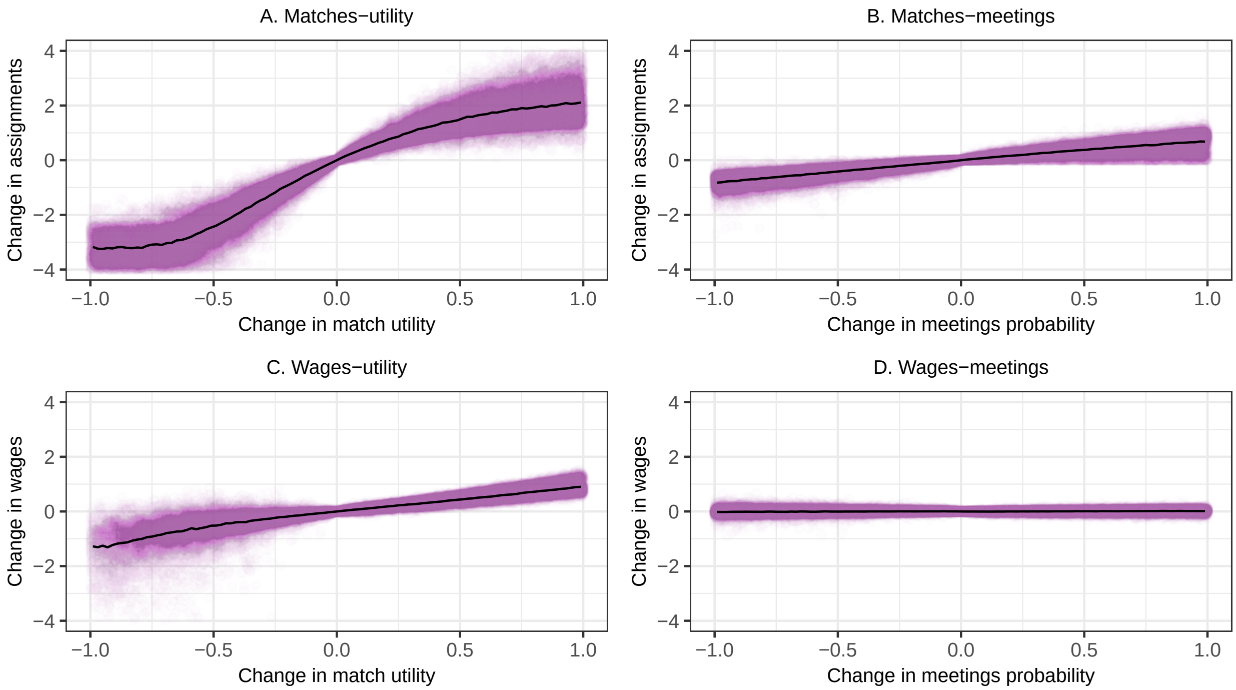
Notes: This figure shows the probability of working in a firm as a function of the difference between the last year the parent's coworker worked at the firm and the worker's labor-market entry year, relative to working in a non-connected firm. The points are the mean coefficients of phantom and weak connections across 100 estimations of equation (5) using a 20 percent random sample of workers each time. I construct the bounds of the 95 percent confidence intervals using the 2.5 and 97.5 percentiles of the coefficients' distribution. The vertical line between -1 and 0 indicates the change from worker-firm pairs with phantom connections to pairs with weak connections.

Figure 3: Event-study plot of coefficients: Effect of weak parental connections on firm assignment, placebo test



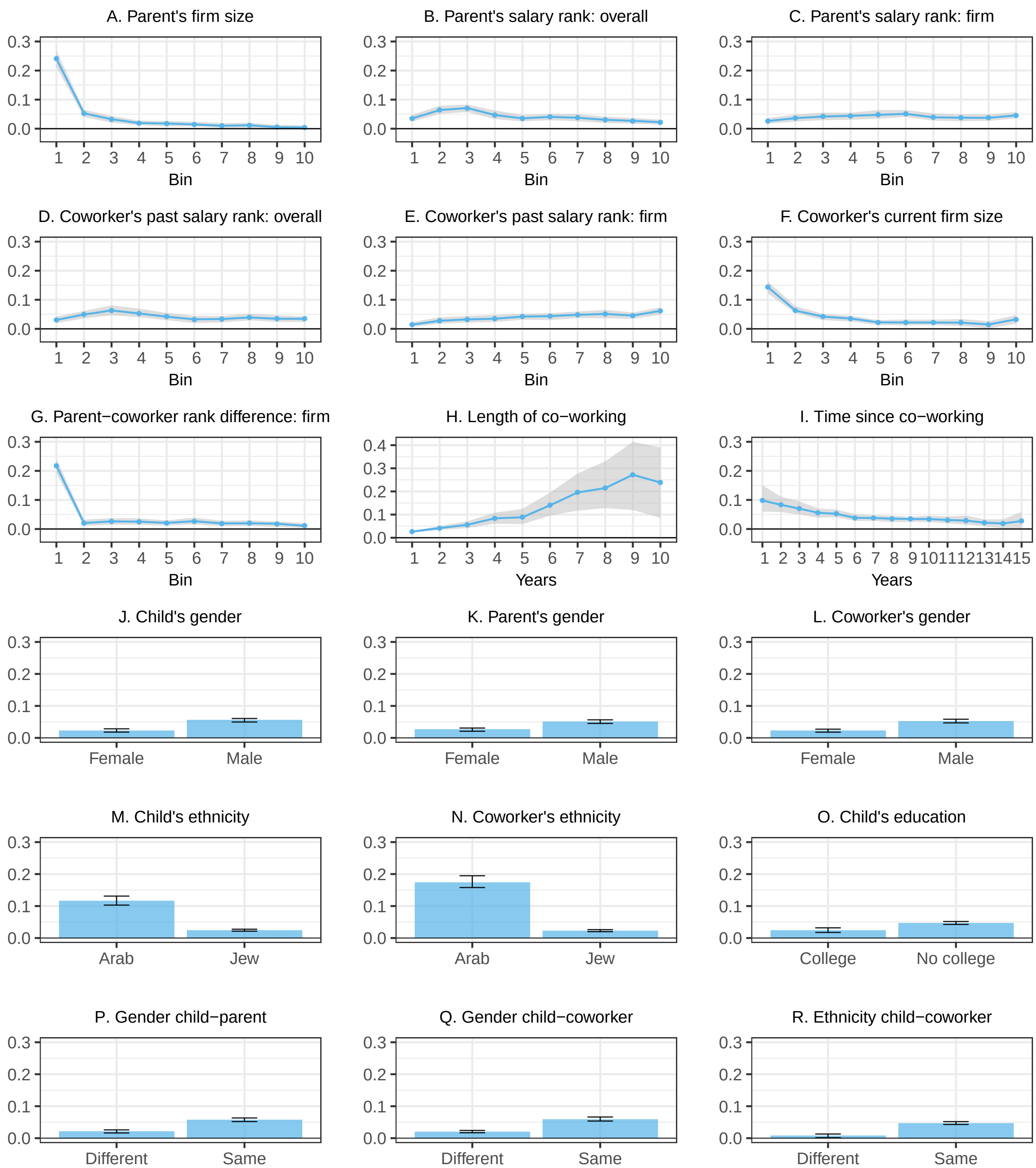
Notes: This figure reports the results of a placebo test, assigning the worker's connections to a random worker in her group. The figure shows the probability of working in a firm as a function of the difference between the last year the parent's coworker worked at the firm and the worker's labor-market entry year, relative to the probability of working in a non-connected firm, based on the new (randomized) data. The points are the mean coefficients of phantom and weak connections across 100 estimations of equation (5) using a 20 percent random sample of workers each time. I construct the bounds of the 95 percent confidence intervals using the 2.5 and 97.5 percentiles of the coefficients' distribution. The vertical line between -1 and 0 indicates the change from worker-firm pairs with phantom connections to pairs with weak connections.

Figure 5: Scatter plot: Changes in outcomes as a result of changes in parameters of the same group of workers and firms



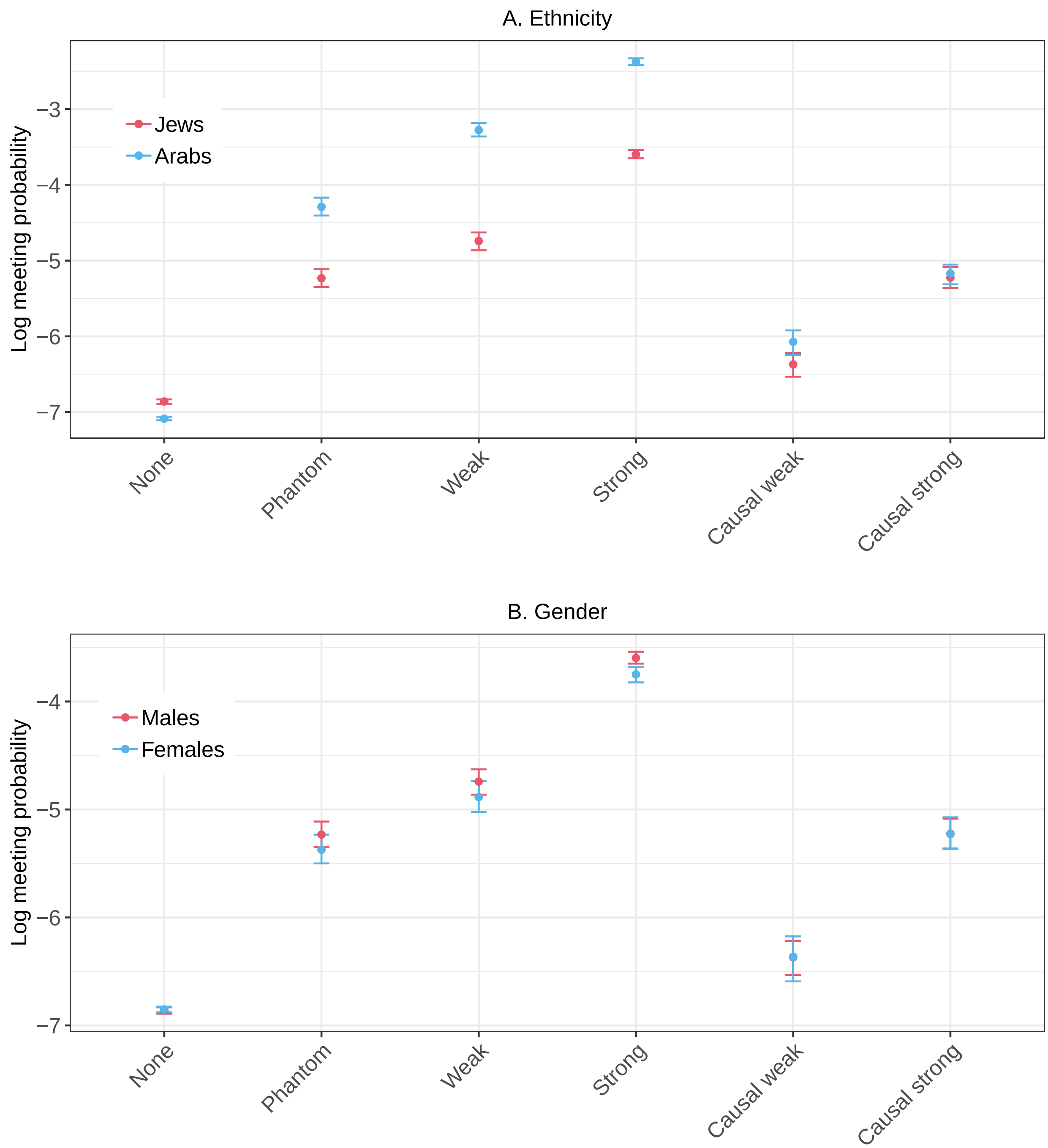
Notes: This figure shows the relationships between the parameters of the model and the predicted moments. I run 10,000 simulations of the model. Each time, I change the value of only one parameter, either the match utility β_{txyc} or the meeting probability p_{txyc}, of one xyc group in each market t by a random number between -1 and 1. Each graph's y-axis is the difference between the (log) number of matches and (log) average wage predicted by the model with the new parameters and the moments predicted with the old parameters. The x-axis is the size of the change to the parameters, β and $log(p)$. The plots show only the results of the moment changes in the $txyc$ cells for which the parameter was changed.

Figure 4: Effects of weak parental connections on firm assignment: Heterogeneity by characteristics of the workers and the connections



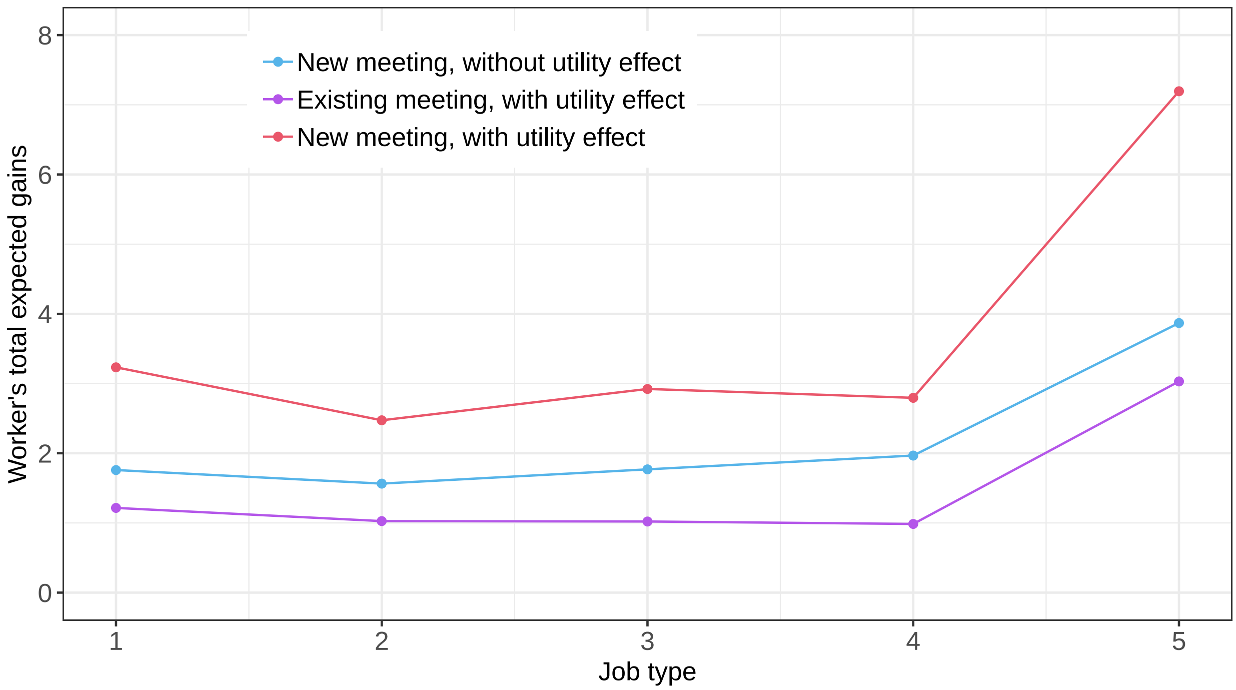
Notes: Each figure shows the probability of working in a firm with weak connections for different characteristics of the workers and the connections, relative to the probability of working in a non-connected firm. The points are the mean coefficients of weak connections across 100 estimations of equation (1) with separate coefficients for different groups of weak connections, using a 20 percent random sample of workers each time. I construct the bounds of the 95 percent confidence intervals using the 2.5 and 97.5 percentiles of that distribution of coefficients.

Figure 6: Model estimates: Average meeting probability by workers' group and connection type



Notes: This figure shows the results of regressing the log of the meeting probabilities obtained from the model on worker, firm, and connection characteristics, and the interactions between worker and connection features. I estimate the regression using weighted least squares, with weights equal to the actual number of matches of the $txyc$ cell. Each point on the graph is the meeting probability by ethnicity and connections type predicted by this regression. Each regression is calculated separately for each of the 100 estimations of the model, and the table reports the averages across the 100 estimations (and their standard errors).

Figure 7: Value of a meeting by job type



Notes: This figure shows the impact of a new meeting or connection on the average worker's expected value separated according to the type of firm with which the meeting/connection is generated. Each line reports the average change in the salary of workers in a different scenario using 5000 simulations: 1) adding a meeting to a random worker and firm in each market, assuming no connections between them, 2) choosing a random non-connected pair in each market and changing the systematic match utility to reflect the utility of a causal weak connection, and 3) adding a random meeting with causal weak connections. The utility of a causal weak connection is the excess utility of weak connections compared to phantom connections.

Appendices

A Appendix Tables and Figures

Table A1: Summary statistics—firms

	1-4	5-500	501+
Firms	123,677	51,999	392
Workers	225,830	1,155,398	833,097
Av. firm size	1.83	22.23	2131.56
Share of firms	0.702	0.296	0.002
Share of workers	0.102	0.522	0.376

Notes: This table reports summary statistics for firms according to the number of workers in the firm. The first row is the overall number of unique firms in 2006-2015 matched employee-employer files. The second row is the total number of workers in each group of firms by year, averaged across the years. The third row is the average number of workers in a firm by year, averaged across the years. The fourth and fifth rows are the share of firms and the share of workers in each group of firms by year, averaged across the years.

Table A2: Balancing test: Correlation between parental connections and measures of proximity between workers and firms

	Log distance	Parent's industry
	(1)	(2)
Phantom connections	-0.369	0.077
	[-0.376,-0.362]	[0.076,0.077]
Weak connections	-0.368	0.076
	[-0.375,-0.361]	[0.075,0.076]
Strong connections	-0.926	0.281
	[-0.944,-0.909]	[0.279,0.284]
R0 (no connections)	10.102	0.033
	[10.090,10.117]	[0.032,0.033]
Ratio weak-phantom	1.000	0.989
	[1.000,1.001]	[0.984,0.995]
Ratio strong-phantom	0.943	2.871
	[0.942,0.944]	[2.850,2.887]
Observations (firms x groups)	21,166,443	21,166,443
N firms	149,729	149,729
N groups	2,959	2,959
N workers	220,684	220,684

Notes: This table reports the (log) geographical distance from the firm and the probability that a firm belongs to the same 3-digit industry of the worker's parent for firms with different types of connections, relative to non-connected firms. The coefficients are the mean coefficients of phantom, weak, and strong connections across 100 estimations of equation (4) with the outcome variables mentioned using a 20 percent random sample of workers each time. I construct the bounds of the 95 percent confidence intervals using the 2.5 and 97.5 percentiles of the coefficients' distribution. R0 is the average outcome variable's value for a non-connected firm. "Ratio weak-phantom" is the estimated odds ratio between the outcome variable's value for a weakly-connected firm and phantom-connected firm. "Ratio strong-phantom" is defined similarly.

Table A3: Effects of parental connections on firm assignment: Robustness to the definition of connection types

	Employment		
	(1)	(2)	(3)
Phantom (single contact)	0.010	0.012	
	[0.009,0.011]	[0.011,0.013]	
Phantom (single + multiple contacts)			0.015
			[0.014,0.016]
Weak (signle contact)	0.050	0.053	
	[0.047,0.054]	[0.049,0.056]	
Weak (single + multiple contacts)			0.095
			[0.091,0.100]
Strong (direct + multiple contacts)	0.487		
	[0.472,0.501]		
Direct		3.091	3.092
		[2.977,3.206]	[2.978,3.207]
Multiple contacts		0.171	
		[0.161,0.181]	
R0 (no connections)	0.005	0.005	0.005
	[0.005,0.005]	[0.005,0.005]	[0.005,0.005]
Observations (firms x groups)	21,166,443	21,166,443	21,166,443
N firms	149,729	149,729	149,729
N groups	2,959	2,959	2,959
N workers	220,684	220,684	220,684
N connections	40,827,833	40,827,833	40,827,833

Notes: This table reports the probability of working in a firm with different types of connections, relative to working in a non-connected firm. The coefficients are the mean coefficients of the different types of connections across 100 estimations of the equivalent of equation (4) using a 20 percent random sample of workers each time. I construct the bounds of the 95 percent confidence intervals using the 2.5 and 97.5 percentiles of the coefficients' distributions. R0 is the average probability of working in a non-connected firm. The first column repeats the baseline specification using three types of connections: phantom connection with a single contact, indirect connection with a single contact ("weak"), and either a direct connection or other types of connection with more than one contact ("strong"). Column 2 estimates a separate coefficient for direct connections and for phantom/indirect connections with multiple contacts. Column 3 combines phantom and indirect connections with one or more contacts.

Table A4: Event-study plot of coefficients: Effect of parental connections on firm assignment

Employment			
Phantom connections		Weak connections	
-5	0.005	0	0.057
	[0.002,0.009]		[0.052,0.063]
-4	0.005	1	0.052
	[0.003,0.008]		[0.043,0.060]
-3	0.007	2	0.042
	[0.004,0.009]		[0.033,0.053]
-2	0.009	3	0.040
	[0.006,0.013]		[0.030,0.053]
-1	0.012	4	0.050
	[0.008,0.015]		[0.035,0.067]
1	0.026	5	0.040
	[0.020,0.032]		[0.032,0.048]
2	0.017	Strong connections	
	[0.013,0.022]		
3	0.013		0.487
	[0.009,0.017]		[0.472,0.501]
4	0.009		
	[0.006,0.014]		
5	0.008		
	[0.005,0.011]		

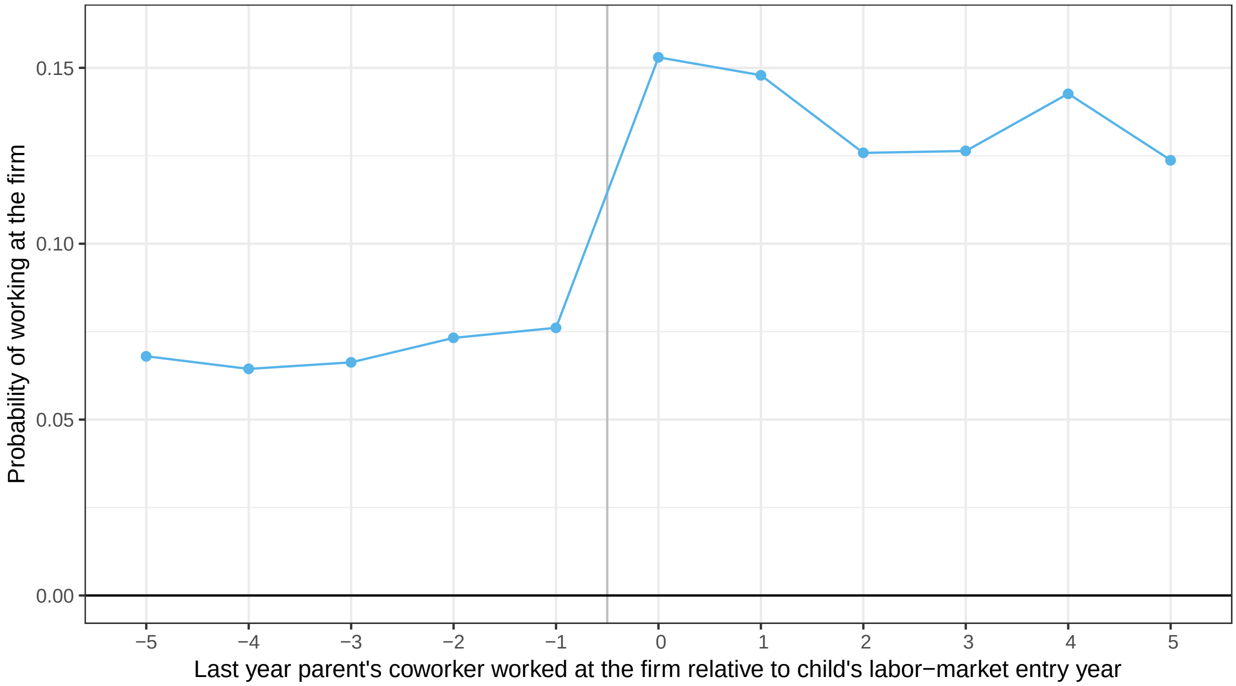
Notes: This table reports the probability of working in a firm for a different types of connections relative to working in a non-connected firm. Phantom and weak connections are divided according to the difference between the last year the parent's coworker worked at the firm and the child's labor-market entry year, The points are the mean coefficients across 100 estimations of equation (5) using a 20 percent random sample of workers each time. I construct the bounds of the 95 percent confidence intervals using the 2.5 and 97.5 percentiles of the coefficients' distribution.

Table A5: Moments-parameters elasticities

	Matches-utility $dln(\mu)/d\beta$ (1)	Matches-meetings $dln(\mu)/dln(p)$ (2)	Wages-utility $dln(w)/d\beta$ (3)	Wages-meetings $dln(w)/dln(p)$ (4)
Same workers and firms	3.511	0.777	3.427	0.015
	(0.078)	(0.017)	(0.325)	(0.009)
Same workers, different firms	-0.264	-0.033	0.001	0.014
	(0.026)	(0.003)	(0.011)	(0.001)
Different workers	-0.008	0.000	-0.032	-0.002
	(0.002)	(0.000)	(0.005)	(0.000)

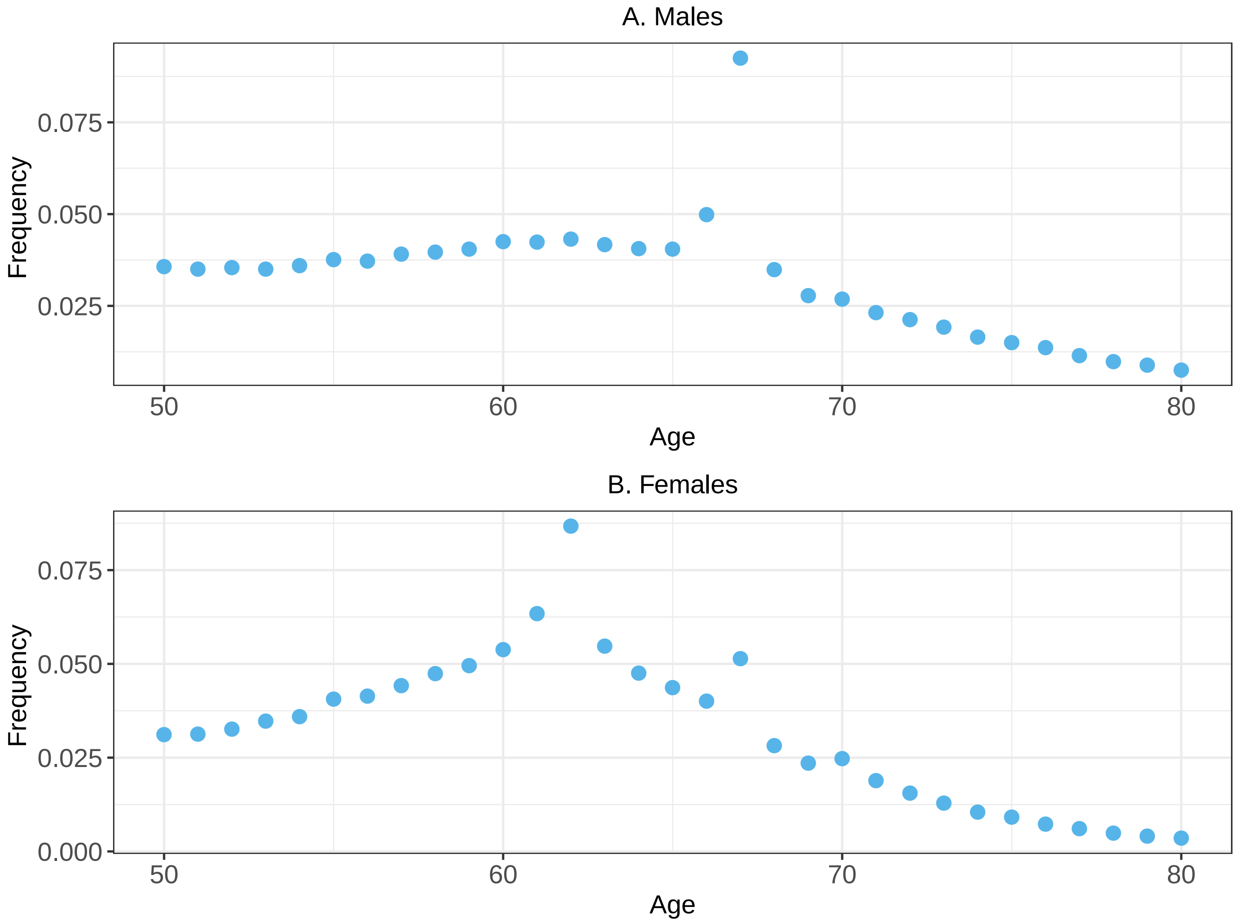
Notes: This table shows the elasticities between the parameters of the model and the predicted moments. I run 10,000 simulations of the model. Each time, I change the value of only one parameter, either the match utility β_{txyc} or the meeting probability p_{txyc}, of one xyc group in each market t by a random number between -1 and 1. Each value in the table is the slope coefficient obtained from regressing the changes in the moment on the parameter changes for different groups of workers and firms. Assume a change in the $txyc$ cell parameters. The first row reports the elasticities of changes in the same $txyc$ cells. The second row reports the elasticities for cells of the type $txy'c'$ where either $y' \neq y$ or $c' \neq c$ (or both). The last row reports the elasticities for cells of the type $tx'y'c'$ where $x' \neq x$.

Figure A1: Raw data: probability of working in a firm for phantom and weak connections



Notes: This figure shows the raw probability of working in a firm as a function of the difference between the last year the parent's coworker worked at the firm and the worker's labor-market entry year. The vertical line between -1 and 0 indicates the change from worker-firm pairs with phantom connections to pairs with weak connections.

Figure A2: Age at last year of work by gender

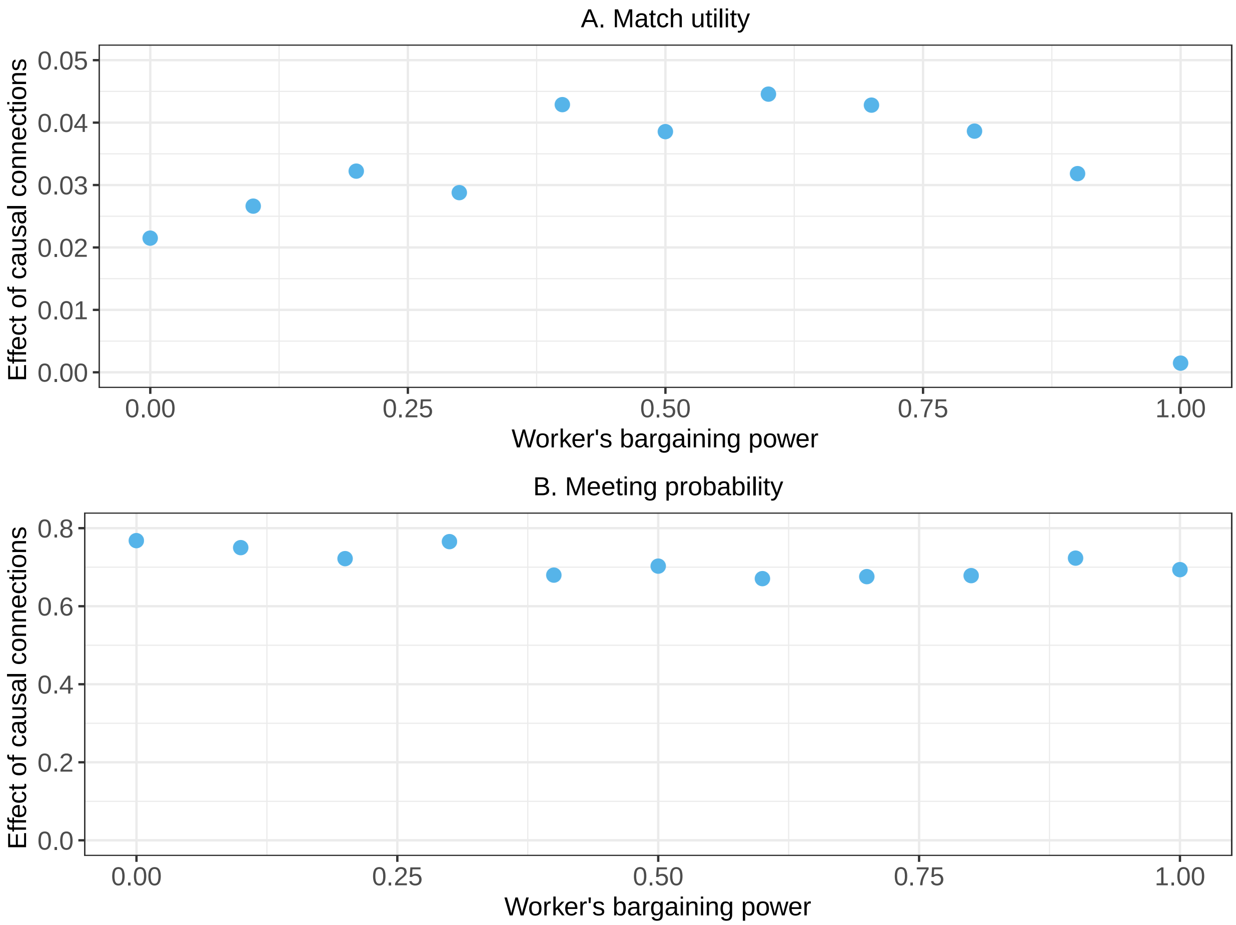


Notes: This figure shows the frequency of the ages of workers when they last appear in the employer-employee data between 2006-2014, separated by gender. Workers that worked in 2015—the final year in the dataset—are not included in this figure. I keep workers that were between 50-80 in their last year of work.

Figure A3: Scatter plot: Lower and upper wage bounds

Notes: This figure shows the relationships between lower and upper wage bounds that support the equilibrium matching. The black line shows the mean value of the wage upper bounds for 100 bins of the lower bounds.

Figure A4: Model estimates of causal weak connections for different values of worker's bargaining power



Notes: This figure shows the model's estimated causal effects of weak connections for the match utility and meeting probability parameters for different workers' bargaining power values. For each worker's bargaining power value, I re-estimate the model and regress the estimated match utility and log of meeting probability parameters on worker, firm, and connection characteristics. I estimate the regression using weighted least squares, with weights equal to the actual number of matches of the $txyc$ cell. Each point on the graph shows the difference between the coefficients of weak and phantom connections for different values of worker bargaining power.

B Data Appendix

This appendix provides additional details on the data preparation and definitions of the variables.

Employment and wages: The data contains observations at the worker $\times$ firm $\times$ year level. For each observation, there are monthly employment indicators and total yearly salaries. In each year, I: 1) drop observations with missing worker or firm identifiers, 2) replace empty monthly indicators with zeros, 3) drop observations that are duplicate in all variables, 4) for duplicate worker-firm observations, take the maximum of the monthly indicators and the sum of the yearly earnings, 5) calculate the monthly salary by dividing the yearly salary by the number of months worked at that firm, 6) keep only observations with employment in February, 7) for each worker, keep the firm with the largest monthly salary 8) keep workers aged 22-69, 9) drop observations with less than 25% of the yearly average wage in the sample.

Education: "No college" workers are defined as workers without any period of enrollment in higher education institutions. Workers with at least one year of admission to higher education institutions (excluding religious schools) are defined as workers with "some college" or simply with "college" education.

Ethnicity: Workers are classified into two categories, Arabs and Jews. Arabs include Arab Muslims, Arab Christians, Druze, and Circassians. In the definition of Jews, I follow the practice of the Israeli Central Bureau of Statistics to include "Jews and Others" together and consider workers without ethnicity classification as Jews.[30]

Workers' location: I measure the new worker's residence location by the longitude and latitude of the city she lived in age 21. I also use the worker's district at age 21, one of seven districts (North, Haifa, Tel-Aviv, Center, Jerusalem, South, and Judea and Samaria).

Natives: Individuals born in Israel and without information on the date of immigration.

Ultra-orthodox: I use the internal algorithm of the National Insurance Institute, which uses information on the residency neighborhoods, educational institutions, and family links to identify Ultra-orthodox individuals.

Industry: I clean the industry variable such that each firm has a unique industry. Using the same employer-employee row file described above, with additional information on the 4-digit industry code of the firm in each observation, In each year, I: 1) drop observations with missing worker, firm, or industry identifiers, 2) for each firm, keep the industry with

[30] According to the Israeli Central Bureau of Statistics definitions, "Others" refer to Non-Arab Christians, members of other religions, and not classified (of Statistics 2019). The vast majority of the people in this category are immigrants from the former Soviet Union who immigrated to Israel in the past three decades. They are not Jews according to the Jewish law but are included in the Law of Return because of their familial ties with Jew (Cohen and Susser 2009).

the most occurrences. Now, if the number of firms in industry A in year t that changed their industry in year $t+1$ to B is greater than the number of firms that stay in industry A, I assume the classification of that industry had changed and update backward industry A to B. Finally, for each firm, I keep the latest industry. In practice, I use the implied 3-digit industry code of each firm.

Firms' location: Unfortunately, exact information on the location of the firms is missing. As a proxy, I calculate the median longitude and latitude of the residence of the workers. I exclude new workers from the calculation of the firms' locations.[31]

Firms' pay premium: I estimate the following AKM model (Abowd et al. 1999)

$$w_{it} = \alpha_i + \psi_{J(it)} + Z'_{it}\gamma + \varepsilon_{it} \tag{B1}$$

where α_i is person fixed effect, $\psi_{J(it)}$ is firm fixed effect, $Z'_{it} =$ are set of year fixed effects and quartic polynomials age-restricted to be flat at age 40 (Card et al. 2018). I estimate the model using workers ages 22-69. I exclude new workers, so their salary would not impact the estimated firm premiums. To capture potential changes in a given firm's premium over the years, I estimate a separate regression each year. Precisely, firm premiums of firms at year t are calculated using the full sample's largest connected set in years $[t-4, t]$. Finally, I rank the estimated firm premiums within a year ("firm rank").

C The role of firms and social networks in earning inequality

In this appendix, I check the raw relationships observed in the data between the ethnicity and gender pay gaps, and firms and measures of the quality of connections.

To get the raw ethnic and gender gaps, I estimate the regression

$$w_i = \gamma_1 \cdot Arab_i + \gamma_2 \cdot Female_i + \phi_{x(i)} + \epsilon_i \tag{C1}$$

using all workers ages 25-65 in Israel in 2015. w_i is the log wage of worker i, $Arab_i$ and $Female_i$ equal 1 if worker i is an Arab or female, respectively. $\phi_{x(i)}$ and $\psi_{j(i)}$ are group and firm fixed effects, respectively. The workers' groups include all combinations of age, education, and district of residence. Columns 1 and 2 of Table C1 report the OLS estimates of equation C1 without and with the firm fixed effects, respectively.

[31] The data do not include an indicator for multi-branch firms. Therefore, I assign the same location for all branches or plants of the same firm. This problem is alleviated by dropping firms with more than 500 workers from the sample.

Starting with the ethnic pay gap, the overall gap between Jews and Arabs in 2015 is 25.3 log points. Controlling for firms, the ethnic pay gap is only 5.1 log points. Comparing the estimates of the ethnic pay gap with and without firm fixed effects, about 80% of the ethnic pay gap in Israel is explained by between-firm variation, and only 20% of the gap is explained by within-firm variation.

The raw gender pay gap, without firm fixed effects, is 36.9 log point. Controlling for the firms the workers work at, the gap decreases to 28.8 log point. Those results indicate that, unlike the ethnic gap, most of the of the gender gap (78%) is explained by within-firm variation.

Table C2, column 1, reports OLS estimates of equation C1 for the sample of new workers. The raw first-job ethnic pay gap is smaller than the population-wide gap (7.7 log points). Controlling for the identity of the firm in which the worker finds her first job, the gap is now positive, where Arabs get 3.0 log points *more* than Jews (column 2).

Column 3 presents a re-estimate of equation C1, including measures of the quality of weak and strong parental connections. The correlation between the average rank of weakly-connected firms and log salary at the first job is positive and statistically significant. The magnitude of the correlation is 1.17 log points per 10 percentile points in the average rank of the connected firms. Interestingly, the magnitude of the correlation is higher for the quality of weak connections than strong connections, with a correlation of 0.90 log points per 10 percentile points in the average rank of connected firms.

Comparing columns 1 and 3 of Table C2, the estimate of the raw ethnic pay gap decreases by about 20 percent when controlling for the measures of parental connections. This result suggests correlational evidence for the importance of parental social connections in the between-group inequality in Israel.

To further explore this, in column 4 of Table C2 I add firm fixed effects to the regression. The coefficients of the correlation between parental connections and salary become very close to zero. Moreover, a comparison between columns 2 and 4 reveals that the estimated within-firm ethnic pay gap is virtually the same, with and without measures of parental connections. Taken together, this suggests that parental social connections are important in explaining the ethnic pay gap in the first job, and only through their impact on the identity of the firm the young workers find for their first job.

To see if the patterns documented for the ethnic pay gap are exceptional, I conduct a similar exercise for the gender pay gap. Table C2, columns 5-8, shows that the gender pay gap patterns are different. Firstly, most of the gender pay gap is explained by within-firm variation (columns 5-6). Secondly, including connections in the regression does not affect the magnitude of the gender pay gap (columns 5 and 7).

In summary, this section suggests that most of the ethnic pay gap in Israel is explained by between-firm variation. Moreover, correlational evidence suggests that better-connected workers find employment at better firms, and that variation in the quality of parental connections explains about 20% of the ethnic pay gap.

Table C1: Ethnicity and gender pay gaps: workers at ages 25-64, 2015

	Log salary	
	(1)	(2)
Arab	-0.253	-0.051
	(0.011)	(0.006)
Female	-0.369	-0.288
	(0.006)	(0.005)
Firm FE	No	Yes
Observations	2,256,441	2,256,441
N firms	188,808	188,808
R^2 (full model)	0.211	0.591
R^2 (projected model)	0.130	0.071

Notes: This table shows the OLS estimates of a wage regression using all workers at ages 25-64 in 2015. The outcome variable is the log of the average monthly wage in 2015. All columns include three dummy variables indicate if the worker is Arab, female, or has some college education, respectively. All columns also include a set of dummy variables for every combination of age, education, and the residential district at age 21. Columns 2 also includes a full set of firm fixed effects. Robust standard errors clustered by group (age-district) and firm are reported in parentheses.

Table C2: Ethnicity and gender pay gaps: new workers

	Log salary			
	(1)	(2)	(3)	(4)
Arab	-0.077	0.030	-0.062	0.030
	(0.004)	(0.003)	(0.004)	(0.003)
Female	-0.203	-0.134	-0.203	-0.134
	(0.003)	(0.002)	(0.003)	(0.002)
Weak con qualiy			0.117	-0.001
			(0.010)	(0.008)
Strong con qualiy			0.090	-0.014
			(0.007)	(0.006)
Firm FE	No	Yes	No	Yes
Observations	211,144	211,144	211,144	211,144
N firms	52,963	52,963	52,963	52,963
R^2 (full model)	0.138	0.614	0.140	0.614
R^2 (projected model)	0.080	0.047	0.083	0.047

Notes: This table shows the OLS estimates of a wage regression using the new-workers sample. The outcome variable is the log of the average monthly wage at the first job. All columns include three dummy variables indicate if the worker is Arab, female, or has some college education, respectively. All columns also include a set of dummy variables for every combination of the year of the first job, age at that year, education, and the residential district at age 21. Columns 2 and 4 also include a full set of firm fixed effects. Finally, columns 3 and 4 include the average rank of the firm pay premiums of the firms that the worker has weak and strong parental connections at. Robust standard errors clustered by group (year-age-district) and firm are reported in parentheses.

D Model Appendix

This appendix provides additional details on the model, its estimation, and the counterfactual exercises.

Moments: Because of limited computational resources in the National Insurance Institute's research laboratory, I perform the estimation of the model outside the laboratory using aggregate information at the level of a $txyc$ cell. For each $txyc$ cell, the information I use is 1) the number of connections d_{txyc}, 2) the number of matches μ_{txyc}, and 3) the average wage w_{txyc}. I calculate the residuals of the wages controlling for groups of year by age and then add the overall mean wage.

To ensure data security, the National Insurance Institute prevent export of any information for groups of less than 10 individuals. Therefore, I do not use matching and wage information on $txyc$ cells with less than 10 matches. In the estimation, I treat these cells as cells with no matches (see below how I deal with such cells). 27.3% of the cells have less than 10 matches, corresponding to less than 1.5% of the workers (and jobs).

Drawing data: I estimate the benchmark model 100 times, each time with a different draw of connections and shocks. Because I cannot use exact information on each worker and firm's connections, I randomly draw d_{txyc} connections of type c between workers of type x and firms of type y at year t. Then, for each worker and firm, I draw random meeting shocks ρ_{ij} from a standard uniform distribution. Likewise, I draw utility shocks ξ_{ij} from a standard normal distribution.

For computational reasons, I keep the information on the shocks of unconnected pairs only if $\rho_{ij} < p_0^{max}$. This is equivalent to the assumption that the meeting probability of unconnected pairs is always smaller than p_0^{max}. I use the value $p_0^{max} = M * T/I$, with $M = 40$, corresponding to an assumption that the average number of meetings per worker with unconnected firms for each txy combination is smaller than 40.

As mentioned earlier, two extra meetings are added to each worker and firm regardless of the model parameters. I do this by setting $\rho_{ij} = 0$ for these pairs.

Normalization: As mentioned in the text, the location of the wages of each market (year) is not determined by the model. I normalize the average wage in each year to the observed mean wage (across all years). I also normalize the meeting probability of the first xyc cell in each market to $M * T/I$, with $M = 20$ meetings per worker on average.[32]

Empty cells: To allow the possibility of $txyc$ cells with no matches, in the estimation equations (23) and (24), I calculate $log(z + 1)$ instead of $log(z)$. Note that in equation (24), the average wage of a cell w_n is multiplied by the number of matches in the cell. Therefore,

[32] Using this normalization, I get an average of 25 meetings per worker (and per job), which is similar to the number of applications per job in Banfi and Villena-Roldan (2019).

there is no need to know the average wage of cells, only the total wage, which allows inclusion of empty cells in the analysis.

Note that because the number of meetings in a cell is bounded below by zero, there is an identification issue in estimating the parameters of empty cells. For example, assume that the model predicts no matches for some $txyc$ cell for a given set of parameters $\theta = (p, \beta, \sigma)$. In this case, decreasing the meeting or utility parameter of this cell will also lead to the same predicted moments. I address this problem in two ways. Firstly, when calculating aggregate statistics and results, such as the average impact of connections on the meeting and utility parameters, I weight each observation by the observed number of matches, which gives no weight to empty cells. Secondly, in the counterfactual exercise, when I calculate the "causal" connection parameters, I cut the top and bottom 1% of outliers, weighted by the number of observations (see more details below).

Negative wages: In principle, the assignment problem can lead to negative values. In practice, after normalizing the average wage in each year to the observed mean wage, I did not get an average negative wage in any of the 100 simulation iterations. If this practical problem does arise, one might use other functional forms instead of the log, such as the Inverse Hyperbolic Sine.

Initial parameter values: To get initial values for the meeting probabilities, I estimate the following equation

$$log(\mu_{txyc}/d_{txyc}) = a + p_c + \epsilon_{txyc} \tag{D1}$$

where d_{txyc} is the share of x-type workers who are c-connected to y-type firms in year t over all possible pairs of x-type workers and y-type firms in year t. Using the weighted least squares estimates (WLS), with weights μ_{txyc}, I calculate $p^0_{txyc} = \bar{p}_0 \cdot \hat{p}_c$.[33]

Similarly, to get initial values for the utility parameters, I estimate the equation

$$log(w_{txyc}) = b + \phi_1 Arab_x + \phi_2 Educ_x + \phi_3 Female_x + \psi_y + \delta_c + \epsilon_{txyc}, \tag{D2}$$

and use the WLS estimates to get the predicted values of each $txyc$ cell. I also use the estimated variance of the error term in that regression for an initial value of σ.

[33]The level of p_{txyc} is not identified together with σ and b. In the estimation, I normalize the first cell in each year to a level $\bar{p}_0 = M \cdot T/I$, where M is the average number of meetings per worker with that level of meeting probability. I choose the value $M = 20$, and use the same level for the meeting probabilities without connections in the initial guess.

Preliminary checks show that the initial values do not have a significant impact on the estimated parameters. I do not systematically explore this direction.

Stopping rule: The algorithm stops when there is no new minimum (lower in ϵ_{tol} from the previous minimum) in the mean squared error of one of the four variables in N_{tol} iterations in a row. I use $\epsilon_{tol} = 10^{-10}$ and $N_{tol} = 50$.

Update rate: In practice, I use $\eta = 0.1$. Using this value, all 100 simulations converged. Once again, I do not systematically explore the conditions for convergence.

Causal connections: The utility parameters of causal connections are calculated as the excess impact of real connections and compared to phantom connections (see equations 28-29). As mentioned above, the estimated accuracy is low for cells with a small number of observations. To account for that, I calculate

$$\beta^{causal}_{txy,weak/strong} = \beta_{txy,weak/strong} + \max\left(\min\left(\beta_{txy,none} - \beta_{txy,phantom}, \beta^{99\%}\right), \beta^{1\%}\right) \quad \text{(D3)}$$

where $\beta^{1\%}$ and $\beta^{99\%}$ are the 1 and 99 percentiles of $\beta_{txy,none} - \beta_{txy,phantom}$, weighted by μ_{txyc}.

Likewise, the utility of a causal connection is

$$p^{causal}_{txy,weak/strong} = p_{txy,weak/strong} \cdot \max\left(\min\left(p_{txy,none}/p_{txy,phantom}, p^{99\%}\right), p^{1\%}\right) \quad \text{(D4)}$$

where $p^{1\%}$ and $p^{99\%}$ are the 1 and 99 percentiles of $p_{txy,none}/p_{txy,phantom}$, weighted by μ_{txyc}.