Who Works Where and Why: The Role of Social Connections in the Labor Market

Shmuel San*

October 3, 2022

Abstract

I develop a two-sided matching model of the labor market with search frictions and use it to study the impact of parental indirect professional connections on the first-job outcomes of children in Israel. Relying on identifying variation from the timing of job movements of parents' coworkers, I find that connections double the probability of meeting and increase by 35% the likelihood of being hired given a meeting. The wage gap between the two major ethnic groups in Israel, Jews and Arabs, decreases by 12% when equalizing the groups' connections but increases by 56% when prohibiting the hiring of connected workers.

*Department of Economics, The Hebrew University of Jerusalem, Jerusalem 91905, Israel. Email: muly.san@mail.huji.ac.il. An earlier version of this article was the first chapter of my dissertation at New York University; thanks to Christopher Flinn, Alfred Galichon, Petra Moser, and Martin Rotemberg for patient advising and support. Thanks also to Milena Almagro, Jaime Arellano-Bover, Barbara Biasi, Pierre-André Chiappori, Christopher Conlon, Oren Danieli, Christian Dustmann, Olivier Godechot, Walker Hanlon, Attila Lindner, Elena Manresa, Konrad Menzel, Maor Milgrom, Fabien Postel-Vinay, Bernard Salanié, Uta Schönberg, Donald Tomaskovic-Devey, Sharon Traiberman, Miguel Uribe, Daniel Waldinger, and audiences at SOLE 2020, NEUDC 2020, EGSC 2020, ESEWM 2020, ZEW 2021, RES 2021, EALE 2021, SIEP 2021, NYU, Hebrew U, Bar Ilan, Tel Aviv, Florida State, Ben-Gurion, Haifa, Collegio Carlo Alberto, IDC Herzliya, Oslo, IIES Stockholm, Helsinki GSE, UCL, and Copenhagen BS for helpful comments and conversations..

1 Introduction

Since the seminal contributions of Shapley and Shubik (1972) and Becker (1973), two-sided matching models with transferable utilities have been used to study a variety of economic questions.[1] In their study of the marriage market, Choo and Siow (2006) suggest to estimate these models by including unobserved heterogeneity in the matching surplus. This allows the model to rationalize real-world data that generally present heterogeneous matching outcomes for observationally equivalent agents.

The contribution of Choo and Siow (2006) is based on three assumptions: agents have perfect knowledge of agents on the other side of the market, there is no sorting on unobserved characteristics on both sides of the market ("separability"), and the latent variables are distributed as iid type I extreme values.[2]

This article develops and estimates a two-sided matching model with transfers of the labor market that relaxes all three assumptions. First, it introduces search frictions into the model, relaxing the perfect information assumption unsuitable in many applications. Second, it develops a simulation-based estimation, allowing for any distribution of the latent variables and their interaction. By that, it relaxes the separability and distributional assumptions.

The transferable utility model seems more appropriate to markets where we know monetary transfers occur in practice, such as the labor market, in contrast to markets like the marriage market. Moreover, this paper shows that by observing the transfers along with the matching outcome, which is common in modern labor-market datasets, it is possible to identify the model's search friction parameters, in addition to the parameters of the matching surplus.

I use the model to study the exact role social connections play in the labor market. The literature offers two main mechanisms for the importance of social connections for matching workers and jobs. First, social connections might reduce search frictions by providing information about job openings at specific firms (to workers) and potential job seekers (to firms). Second, conditional on that mutual knowledge, social connections may increase the probability of a match (a hire) between a job seeker and a firm.

Specifically, the model assumes that 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 the 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 (that might also be affected by social connections) and the set of worker-specific wages that clear the market. To separately

[1] See the references in Galichon and Salanié (2020).

[2] Galichon and Salanié (2020) relax the third assumption regarding the distribution of the latent variables. Nevertheless, they maintain the first two assumptions of perfect information and separability.

identify the two mechanisms, I use two distinct types of information: where individuals end up working (the matching) and how much they are paid (the transfers).

I estimate the model using a novel simulation-based method that allows non-parametric estimation of the meeting rates and matching surplus along with rich and flexible functions of the latent variables. Finding the model's equilibrium matches and wages is computationally feasible due to the sparsity of the choice problem resulting from the model's first stage, which restricts the set of potential matches. I develop a method to estimate the model with a high-dimensional parameter space using an update mapping that "inverts" the information on the observed matches and wages into the meeting probability and match-surplus parameters.

I apply the model to study the impact of the professional network of parents on the first-job outcomes of children in Israel. I begin by distinguishing between strong and weak parental connections. Strong (direct) connections are connections between employees and firms where their parents have worked. Weak (indirect) connections are between employees and firms where their parents' past coworkers have worked. Despite the importance of workers' network of past coworkers for their labor market outcomes (Cingano and Rosolia 2012; Eliason et al. 2022), and the fact that, for each worker, there are many more firms with weak parental connections than with strong ones, there is no research on the impact of the network of parents' past coworkers, which is the focus of this study.

A naive comparison between connected and unconnected worker-firm pairs might attribute the effect of omitted variables such as geographical distance and industrial similarity to the estimated impact of connections. For example, a worker might be more likely to work for companies in her parent's industry, regardless of social connections, since both she and her parent possess similar skills. To identify the effect of weak connections (both in the reduced-form and structural estimation), I leverage the timing of 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 with *non-active* links, that is, where the contact had left a short time before or had joined a short time after the labor-market entry year ("phantom connections"). I show that firms with weak and phantom connections are similar on a variety of characteristics such as sector and location.

Reduced-form estimates show that workers are 3 to 4 times more likely to find employment in firms with (active) weak parental connections than in phantom-connected firms. Workers' probability of starting at a particular firm discretely falls the year after the link is destroyed.[3] Connections are more effective if formed at smaller firms, for more extended

[3]To check for the possibility that estimated effects reflect endogenous separations, I also estimate the regression using two exogenous causes of separation; coworkers' deaths and retirements. These estimates are

periods, and more recently. Notably, connections are also stronger if the child, parent, and parent's coworker share characteristics such as gender or ethnicity.[4] Likewise, the effects are larger for males, from the Arab minority, and less-educated workers, as well as during high unemployment years.

Reduced-form estimates also show that weak connections are associated with 1.4 to 2.5 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.[5] The structural model addresses this issue and other limitations of the reduced-form estimation by jointly studying questions of matching and wage-setting.[6]

The model's estimates suggest that both the "search frictions" and "match surplus" mechanisms are important in explaining why parental connections increase the probability of working in a firm. Weak connections increase the meeting probability by 115% and the likelihood of being hired given a meeting by 35%.

To study the wage effects of connections, I evaluate two sets of counterfactuals. Both counterfactuals rely on the assumption that the connections' causal impact (or the impact of "causal connections") is the excess effect of real connections relative to phantom connections. In the first set of counterfactuals, I evaluate the wage-equivalent value of meetings and connections. I find that the average value of one additional meeting with an unconnected firm is 2.2% of the new workers' average wage. 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% of 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 this effect is due to workers moving to the new connected firm, whereas the remaining 16% is due to improving workers' choice set without changing their job.

The model can be used to examine how parental connections affect wage inequality between groups. Specifically, 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

similar in magnitude to the benchmark result. Likewise, to check the potential difference in employment trends in firms with 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.

[4] That is to say, for example, that fathers' connections matter more for boys and mothers' for girls.

[5] 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 firm hires the worker.

[6] The reduced-form estimation abstracts from spillovers and equilibrium effects. The model addresses it by considering the full structure of connections in the economy in an equilibrium framework. See the beginning of Section 4 for further discussion.

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.

This paper contributes to several strands of the economics literature. First, it contributes to a body of research that studies the effects of parental connections on labor-market outcomes. Existing literature finds that direct links (where parents work) increase the child's probability of working there (Corak and Piraino 2011; Kramarz and Skans 2014; Stinson and Wignall 2018; Staiger 2021; Eliason et al. 2022); however, there is less evidence for the impact of indirect parental connections. Existing studies find no impact for weak or indirect parental connections, such as parents of high-school classmates (Kramarz and Skans 2014) or high-school classmates of one's parents (Plug et al. 2018). The positive effect I find for the channel of parent's past coworkers' network compared to other channels of indirect parental networks is consistent with a literature showing the importance of coworker networks for worker's own labor market outcomes (Granovetter 1973; Cingano and Rosolia 2012; Hensvik and Skans 2016; Caldwell and Harmon 2019; Eliason et al. 2022).[7]

This paper offers a new identification strategy for the effect of indirect parental connections on labor market outcomes: comparing active and non-active links (when workers made the employment decision). Compared to looking only at parents' employment, studying the entire network of parents' coworkers provides more useful variation. Moreover, the assumption that the timing of job movements of contacts is orthogonal to the workers' labor market entry makes much less sense if applied to the parents themselves rather than the parents' coworkers.

Second, this paper adds to the understanding of the mechanisms for which social connections are helpful in matching workers with firms. As discussed earlier, the literature points to two main mechanisms. First, social connections could facilitate the flow of information regarding job opportunities and job seekers (Calvo-Armengol and Jackson 2004; Fontaine 2008). Second, connections might impact the value of the prospective match, either by affecting the match's productivity (Bandiera et al. 2009), favoritism (Beaman and Magruder 2012; Dickinson et al. 2018), or by reducing uncertainty regarding the productivity of the worker or the match (Montgomery 1991; Dustmann et al. 2016; Bolte et al. 2020).

[7] I also find that the effect of connections decays over time, which explains why links formed a long time ago are not useful.

In this paper, I build and estimate a matching model that separately identifies these two mechanisms. To the best of my knowledge, this is the first work that studies these mechanisms in a joint framework. Differentiating between these two mechanisms is essential for predicting the effectiveness of different policy measures. For example, if the second mechanism is the one that matters, then merely encouraging job interviews is unlikely to have a sizable impact. In contrast, other policies, such as subsidizing long-term internships, are likely to have an impact through both mechanisms.

Third, I contribute to the two-sided matching literature by introducing search frictions into this type of model and allowing any distribution type of the latent variables. Up until now, those models assumed that each agent has perfect information about all agents on the other side of the market and can choose each one of them (Choo and Siow 2006; Chiappori and Salanié 2016). My model departs from the perfect information assumption by restricting the feasible choice set of the agents. This extension empowers the model to study markets where search frictions are important, such as the labor market. I exploit the assignment problem's sparsity implied by the search-frictions assumption, together with recent developments in assignment problem algorithms, to simulate the model. Thus, I can estimate the model using simulations even with large-scale data, allowing non-parametric systematic surpluses and any parametric assumption regarding the distribution of the latent variables. In particular, the model relaxes the separability assumption which is in use in the vast majority of this literature (Salanié 2015; Chiappori et al. 2017; Galichon and Salanié 2020).[8]

Fourth, I contribute to the literature that models search frictions in the labor market. The model in this paper offers a new "technology" to model search frictions in a static framework, using a restricted choice set that the two sides can choose from.[9] Compared to random search models, where workers meet firms sequentially (Mortensen and Pissarides 1994; Burdett and Mortensen 1998; Shimer and Smith 2000; Postel–Vinay and Robin 2002; Hagedorn et al. 2017), the proposed model can better answer questions regarding the quality of jobs people find, in which the important margin is not necessarily binary (whether to work/switch jobs).

[8] See Fox et al. (2018) for a notable exception. See also Agarwal (2015) for a 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 Del Boca et al. (2014) for a matching model (of the marriage market) with restricted choice sets. Finally, see Caldwell and Danieli (2021) for a recent study that uses a two-sided matching model (with perfect information) to derive a sufficient statistic for studying the effect of outside options on wages.

[9] A potential future dynamic version of the model will enable the study of many more important issues, such as learning of the match quality over time or dynamic considerations of working in a firm due to on-the-job social network and human capital formation, to mention but a few (see Dustmann et al. (2016), Bonhomme et al. (2019), and Arellano-Bover and Saltiel (2021).) However, unlike standard search and matching models, the dynamic is not required here to model the search process itself.

However, unlike directed search models (Shi 2002; Menzio and Shi 2011), this model leaves room for factors such as social connections to provide new information about the existence of vacancies and candidates.[10]

Fifth, this paper contributes to the literature that studies the importance of social connections not only for individuals but for society at large, and particularly for income inequality. Theoretically, due to the homophily in social networks (McPherson et al. 2001), workers from advantaged groups have better connections, and therefore social connections will further increase the initial between-group inequality (Calvo-Armengol and Jackson 2004; Bolte et al. 2020). Empirically, a recent paper by Miller and Schmutte (2021) shows that, indeed, referral hiring helps to explain racial differences in various labor-market outcomes in Brazil.

Using estimates from my model, I examine how social networks contribute to pay gaps between Arabs and Jews in Israel. Indeed, a non-trivial part of the gap between the groups can be explained by differences in the quality and quantity of connections, as the theory suggests. Nevertheless, my results also indicate that, in total, hiring through social connections *reduces* between-group inequality since the disadvantaged group uses it more extensively.[11]

The rest of the paper proceeds as follows. Section 2 describes the data, definitions, and identification strategy. Section 3 presents the reduced form framework and its results. Section 4 develops the model and the estimation method. Section 5 quantifies the impact of parental connections, Section 6 evaluates the wage effects of connections using counterfactual analysis, and Section 7 concludes.

2 Data and identification strategy

2.1 Data and definitions

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 employment indicators for each firm in which a person worked, the yearly salary received from each firm in a 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

[10] A comprehensive comparison of the proposed model to existing search and matching models of the labor market is outside the scope of this paper.

[11] Kramarz and Skans (2014) also find that the effect of parental ties is stronger for young workers with less education, lower GPA grades, and generally with poor labor market prospects. However, they do not explore the inequality consequences of these differences.

birth, date of death (if any), sex, ethnic group, country of birth, and date of immigration to Israel. Most important for this study, the data include identifiers of the parents of each individual, which enables me to link parents and children.

This paper studies the impact of the professional network of parents on the employment and salary of young workers entering the labor market. The paper's primary focus is on weak or indirect parental connections, which are the connections between workers and firms in which exactly one of their parent's past coworkers work at their labor market entry year. For comparison, I also study the effect of strong or direct connections, which satisfy at least one of the following conditions: 1) the worker's parent worked at the firm in the past, 2) more than one of the worker's parent's past coworkers worked at the firm at any time within five years before or after the worker's labor market entry year.[12]

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 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 stable job.[13]

My analysis sample comprises Israelis who found their first stable job between ages 22-27 in the years 2006-2015 in a 5-500 workers firm. I exclude workers without any parent that worked in a 5-500 workers firm when they were 12-21 years old. I further exclude immigrants and Ultraorthodox Jews from the sample.[14]

2.2 Summary statistics

Table 1 shows sample sizes and sample means. The new workers' sample—my main analysis sample—includes 220,806 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 5,839 NIS, which is equivalent to 1,621 USD (2017 prices).

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

[12]To correctly measure the impact of weak connections, in the main analysis, I define both the treatment (weak connections) and control (phantom connections, see below) groups as firms in which precisely one of their parent's past coworkers worked. I allocate firms with more than one connection in the "strong" connection group. However, I also check the robustness of the results for alternative definitions of connections (Table A4).

[13]I check the robustness of the results for two alternative definitions of the labor-market entry year : 1) The year of which the worker is 25 years old, and 2) The year after the worker's graduation year (Figure A3).

[14]See Appendix B for further information about the data construction and the definitions of the variables.

who find their first job in a connected firm is higher for Arabs than for Jews (Table 1).

Comparing males and females, males earn more at their first job but work at similar-paying firms to females. Likewise, males are connected to firms with slightly lower rank (in terms of AKM pay premium) than females.[15] 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 five bins using their 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).[16]

2.3 Identification strategy: comparing real and phantom connections

How much more likely the average worker is to work in a connected firm than in an unconnected firm? A naive 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 the impact of 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 with specific firms and impact the employment probability.

[15]This difference is due to the fact that the participation rate of Arab women in the labor market is much lower than that of Jewish women.

[16]See appendix C for the correlation between the ethnicity and gender pay gaps on the one hand, and firm pay premiums 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 ("weak connections") with a firm with *non-active* social connections ("phantom connections"). In particular, it compares the likelihood of working for a firm where one of the worker's parents' past coworkers worked during the labor market entry year with the likelihood of working for a firm where the contact had worked there within five years before or after the worker entered the labor market, but not that year.

The identification strategy suggested relies on the assumption that the movements of the parents' past coworkers between firms are orthogonal to the job decision of the children. Under that assumption, there are no systematic differences between firms with active and non-active connections, except for the existence of the contact at the appropriate time at the actively (weakly) connected firm. Therefore, comparing the two types of firms will isolate the impact of connections themselves.[17]

Note that phantom connections might still impact the probability of working in a firm. The past employer of a specific firm might deliver relevant information to her contact, either because of the past knowledge she has about the firm or the current links she still has in the firm. Therefore, the estimates obtained using this identification strategy are lower bounds of the actual effect.[18]

Compared to looking only at parents' employment, studying the entire network of parents' coworkers provides more useful variation. Moreover, the assumption that the timing of job movements of contacts is orthogonal to the workers' labor market entry makes much less sense if applied to the parents themselves rather than the parents' coworkers.[19] For example, the parent might have left the firm because the family had moved to a different location, which might by itself reduce the probability of the child working at that firm. However, it is harder to think of similar threats related to the timing of the coworker movements.

[17]This comparison also addresses the potential measurement error problem using coworkers to define connections, as it occurs symmetrically both in the treatment and control groups.

[18]Alternatively, another type of "phantom" connection might be workers that worked at the same past firm as the parents, but in different years; see Hensvik and Skans (2016) for a similar approach. However, using that type of connection as the control group assumes that the (past) job movements of the parents are orthogonal to the job decision of the children, which is less plausible in the context of this paper.

[19]Kramarz and Skans (2014) compare situations when the parent is present in the actual firm versus when the parent has recently left the firm in some of their robustness tests.

3 Regression results

3.1 Regression framework

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.

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

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

e_{ixj} is an indicator variable taking the value one if individual i from a group x starts working in firm j. ϕ_{xj} is a match-specific effect that captures the propensity that a worker from a given observable group ends up working in a particular firm. Workers' groups (x) include all combinations of ethnicity, gender, education, age, year of the first job, and district of residence of the new workers. D_{ij}^p, D_{ij}^w, and D_{ij}^s 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 estimate 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. Comparing the impact of weak and phantom connections, allows me to isolate the effect of weak connections from other factors that might be correlated with them.[20]

Unlike Kramarz and Skans (2014), I do not assume that $E(\epsilon_{ixj}|D_{ij}^s, D_{ij}^w, D_{ij}^p, x \times j) = 0$. This assumption is not plausible as jobs with some sort of connections are different from jobs without any connections.[21] Instead, my identification strategy assumes that $E(\epsilon_{ixj}|D_{ij}^w, x \times j) = E(\epsilon_{ixj}|D_{ij}^p, x \times j)$. Using that assumption, I can identify $\delta^w - \delta^p$. Sections 3.3 and 3.4 provide evidence to support this identification assumption.

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 an extended version of the fixed-effects transformation,

[20] This specification is abstract from spillovers and equilibrium effects. For example, the probability of working at a firm j might also depend on the probability of working at any other firm j', which in turn will depend on the connections to j'. The model in next the section explicitly addresses these issues.

[21] For example, jobs with one of three types of connections are geographically closer and more likely to belong to the parent's industry relative to non-connected jobs (Table A6).

proposed by Kramarz and Thesmar (2013) and Kramarz and Skans (2014).[22]

3.2 Regression results

Table 2 presents estimates of the coefficients in equation (1).[23] 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, implying 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 than 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-connected firms. The estimated probability of working in a weakly- (strongly)-connected firm is 3.7 (32.8) times higher than the probability of working in a phantom-connected firm for the average new worker (Table 2, column 1).[24]

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 4.19 times higher than a phantom-connected firm for Arabs and 3.31 times for Jews. Similarly, the effect of weak connections was stronger for males (3.96) than females (2.97) (Table 2, columns 4-5).[25]

Overall, the findings here about positive and large impact of strong connections are consistent with existing literature (Corak and Piraino 2011; Kramarz and Skans 2014; Staiger 2021). However, existing literature finds no impact for weak or indirect parental connections, such as parents of high-school classmates or high-school classmates of one's parents (Kramarz

[22]See Appendix D.1 for further information about this methodology.

[23]To ease visualization, I scale the employment outcome by 100 in all the employment specifications. Hence, the results are in terms of percentage points.

[24]The weak-phantom ratio is calculated as follows. The unconditional average probability of working in a non-connected firm (R0) is 0.005 percentage points. Therefore, from equation (1), the estimated (average) probability of working in a weakly connected firm is $0.005 + 0.050 = 0.055$ percentage points. Likewise, the estimated (average) probability of working in a phantom connected firm is $0.005 + 0.010 = 0.015$ percentage points. The ratio between the two probabilities is $0.055/0.015 = 3.7$. The strong-phantom ratio is calculated similarly.

[25]See Section 3.5 below for a further discussion about the heterogeneous effect for different groups of workers. See also Appendix D.2 for the robustness of the results for alternative definitions of connections and labor-market entry year.

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 parental networks is consistent with a literature showing the importance of coworker networks for worker's own labor market outcomes (Granovetter 1973; Cingano and Rosolia 2012; Hensvik and Skans 2016; Caldwell and Harmon 2019; Eliason et al. 2022).[26]

3.3 Event study

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

$$e_{ixj} = \phi_{xj} + \sum_{c=p,w} \sum_{\tau=-5}^{5} \delta^{c,\tau} \cdot D_{ij}^{c,\tau} + \delta^{s} D_{ij}^{s} + \epsilon_{ixj} \tag{2}$$

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.[27]

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.

The estimates of the coefficients in equation (2) 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.[28]

The probability that a new worker began work at a firm that her parental contact left

[26] See also Section 3.5 below, where I show that the effect of connections decays over time, which might explain why links formed a long time ago are not useful.

[27] Note that, for $\tau < 0$, the contact left the firm before time zero (the labor-market entry year), therefore $D_{ij}^{w,\tau<0} = 0 \ \forall i, j$. Similarly, if i's contact left the firm at time zero, i cannot have phantom connections at that firm: $D_{ij}^{p,\tau=0} = 0 \ \forall i, j$.

[28] 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 A5 reports all estimated coefficients of equation (2). 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 are significantly smaller than the parallel effect for weak connections.

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 0.040-0.058 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.[29]

3.4 Validating the identification strategy

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. The event-study results presented above support this assumption. In Appendix D, I support this assumption in three additional ways. First, I show that firms with weak and phantom connections are similar on observable characteristics (Appendix D.3).

Second, 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.[30] These estimates are similar in magnitude to the benchmark result, with odds ratios of 2.6 and 3.9 for the "death" and "retirement" connections, respectively (Appendix D.4).

Third, to check the potential difference in employment trends in firms with 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 (Appendix D.5).

3.5 Heterogeneity of the effect

Is the impact of parental connections on employment similar for workers with different characteristics? How do the characteristics of the connections themselves change the effect? To check the heterogeneity of the effect, I re-estimate equation (1) with separate coefficients for different categories of weak and phantom connections. Figure 3 shows the difference between the estimates of the effect of weak and phantom connections on employment for

[29]Similar patterns are obtained when estimating equation 2 for sub-groups of workers divided by ethnicity and gender (Figure A2).

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

each category. Below are the main findings.[31]

The quality of connections. Connections that formed at smaller firms are more effective (Figure 3, 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 B). This fact also can be explained by a higher probability that the contact can impact the hiring decisions in smaller firms.

Next, the effect is more substantial the longer the parent and the contact worked together. Likewise, the effect is weaker for connections generated less recently (Panels C and D). Overall, these results are consistent with the results of Eliason et al. (2022), who find that the impact of social connections is stronger for connections of longer duration, more recently established, or fostered in smaller groups.

Parent's and coworker's salary. I check the magnitude of the effect as a function of the overall (countrywide) and within-firm salary rank of the parents and coworkers. Starting with the overall salary rank, Panels E and F of Figure 3 show that, except for the two lowest wage percentiles, the effect is smaller the higher the salary of the parents and the coworkers, indicating that workers from disadvantaged backgrounds use connections more. On the contrary, the effect tends to increase with the relative salaries of parents and coworkers in the firm (Panels G and H).[32] Moreover, the smaller the wage gap between the parent and the coworker, the stronger the effect (Panel I).

Gender. The effect is stronger for males than for females. This fact is true when considering the gender of the worker, the parent, and the parent's coworker (Figure 3, Panels J-L). This result is in line with Bayer et al. (2008) and Kramarz and Skans (2014) who show that social networks are less effective for females.

Ethnicity and education. The effect is stronger for Arabs than Jews and weaker for more highly-educated workers (Panels M-O). It is consistent with the results above that the effect is stronger for workers from disadvantaged backgrounds.[33]

Similarity between the child, the parent, and the coworker. The effect is stronger if the parent, the worker, or the parent's coworker are of the same gender (Panels P-Q). Likewise, the effect is stronger if the worker and the parent's coworker are from the same

[31] Note that when comparing between the effects of two groups, one cannot distinguish between the two following alternatives without direct information on the actual connections between the workers: 1) the probability of connection formation is higher for one group compared to another, and 2) the impact of the connections is higher. Therefore, caution is needed when interpreting the estimates.

[32] The fact that links to higher-ranking employees within a firm are more effective is consistent with Montgomery (1991)'s model. See Kramarz and Skans (2014) for similar results.

[33] Kramarz and Skans (2014) also find that the effect of parental ties is stronger for young workers with less education, lower GPA grades, and generally with poor labor market prospects.

ethnic group (Panel R).

Unemployment rate. Figure A6 shows the correlation between the estimated effect of weak connections by year and the total unemployment rate in that year. The effect is stronger in years with high unemployment rate, in line with the results Kramarz and Skans (2014) find for strong parental connections.

3.6 Correlation with salary, job tenure, and salary growth

Next, I turn to check the relationships between parental links and other labor-market outcomes of new workers. Specifically, I compare the wage and tenure at first job and the salary growth after three years of workers from the same observable group, with and without connections to the firm where they found their first job. To account for factors correlated with parental connections, I compare real and phantom connections. I also add firm fixed effects in part of the specifications.

Precisely, I assume that the labor-market outcome y_i of a new worker i equals:

$$y_i = \sum_{c=p,w,s} \delta^c D^c_{i,j(i)} + \phi_{x(i)} + \psi_{j(i)} + \epsilon_i. \tag{3}$$

where $D^c_{i,j(i)}$ 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 the first job, and district of residence of the new workers.

The first two columns of Table 3 report the estimates of equation (3) with log salary as the outcome variable, with and without firm fixed effects. Without controlling for the firm in which the workers found their first job, the salary of workers with phantom connections is lower by 0.7 log points than observably similar workers without connections (not statistically significant). However, having real connections at the firm, either weak or strong, is correlated with a higher salary than workers without connections. The coefficients are 1.8 and 7.4 log points for weak and strong connections, respectively. Compared to phantom connections, weakly and strongly connected workers' salaries were higher by 2.5 and 8.3 log points (Table 3, Column 1)

Column 2 of Table 3 shows the estimates with firm fixed effects. The salary of workers with phantom, weak, and strong connections to the firm is higher by 1.2, 2.6, and 8.3 log points than observably similar workers at the same firm without connections. Compared to phantom connections, weakly and strongly connected workers' salaries were higher by 1.4

and 7.1 log points.

The third and fourth columns of Table 3 investigate whether workers with a connection at their first firm stay at that firm for more extended periods than unconnected workers. The outcome variable in columns 3 and 4 is the number of years the worker stayed at her first firm. Without (with) controlling for firm fixed effects, the first-job duration of workers with phantom, weak, and strong connections is higher by 0.123 (0.098), 0.182 (0.187), and 0.601 (0.441) years, respectively, compared to workers without connections. Compared to phantom links, weak and strong connections are correlated with 0.059 (0.089) and 0.419 (0.343) more years at their first firm.

The last two columns of Table 3 compare the salary growth after three years (not necessarily at the same job) for workers who started their first job at a connected and unconnected firm. Without (with) controlling for firm fixed effects, the salary growth of a worker who started her first job at a weakly connected firm is 4.5 (2.7) percents lower compare to phantom connections. However, there is no significant difference between the wage growth of strong and phantom connections.

What can explain the results that connected workers receive higher entry salaries, are less likely to leave the firm, and experience a slower wage growth? Jovanovic (1979) shows that if the new entrants' productivity is very noisy, the firm will offer a low initial wage, high wage growth, and the worker's turnover rate will be high (as low productive workers will exit soon). The higher entry wages and tenure and slower wage growth (for weak connections) found here are consistent with that explanation if connections reveal information about the new entrant's skill level. Other causes that link hiring through social connections and higher firm utility are also consistent with my results.

My results are also consistent with the results of Dustmann et al. (2016), who find that workers that are more likely to find their jobs through referrals receive higher entry wages, experience slower wage growth, and are less likely to leave the firm. However, the results are inconsistent with Kramarz and Skans (2014) who find lower starting wages and higher wage growth for entrants hired through (strong) parental connections.

Comparing worker-firm pairs with real and phantom connections helps isolate the relationships between these outcomes and social connections from other factors correlated with connections, such as geographical distance and industrial similarity. However, because connections also impact the identity of the firms the workers end up working at (and for which we observe the wage information), naive wage regressions cannot identify the causal impact of connections on wages. 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 surplus for different worker-firm matches. Likewise, although not explicitly modeled, the correlation between parental connections and job duration is consistent with my finding of higher match surplus the firms get from hiring connected workers. I discuss these issues in more detail below.

4 A two-sided matching model of the labor market with transferable utilities and search frictions

Social connections are valuable for workers entering the labor market for two main reasons. First, they might alleviate search frictions by improving the information flow about a job opening at a specific firm and a potential job seeker. Second, conditional on that mutual knowledge, they may increase the probability of a match between the job seeker and the firm.

In what follows, I evaluate the role of the two mechanisms by building and estimating a two-sided matching model of the labor market with search frictions. Typically, the two-sided matching literature assumes that each agent has perfect information about all agents on the other side of the market and can choose each one of them (Choo and Siow 2006; Chiappori and Salanié 2016). In my model, I depart from the perfect information assumption by restricting the feasible choice set of the agents.

Precisely, I assume that 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, which might also be affected by social connections.

Using this conceptual framework, I separate the potential mechanisms offered in the literature for the importance of social connections for matching workers and firms into two groups. In the first group of mechanisms, social connections reduce job search frictions by improving the information flow about open vacancies and potential candidates. In the second group, connections directly impact the value of the prospective match. To the best of my knowledge, this is the first model that combines the two types of mechanisms usually studied in isolation into a joint framework.[34]

Disentangling the two mechanisms described above is essential to predict the effectiveness of different policy measures. For example, suppose connections are valuable mainly because they alleviate search frictions. In that case, policies that aim to create more job interviews

[34]See Appendix E.1 for further discussion about the links between the current model and existing theoretical models of social connections.

between workers from disadvantaged groups and high-paying firms are likely to substitute social links.[35] However, it is less plausible to assume that such policies can generate additional value to the match. Therefore, if the "match surplus" channel is dominant, the effectiveness of such policies will be more moderate.

The structural estimation of the model also allows the evaluation of counterfactuals accounting for spillovers and equilibrium effects. Using the reduced-form estimates to do so might lead to bias conclusions for at least three reasons. First, the reduced-form estimation implicitly assumes no spillovers between workers and between firms. In reality, however, the probability of worker i of working at a firm j might depend on the probability of other workers working at that firm, which in turn will depend on the connections they have in the firm. Likewise, the probability of worker i of working at a firm j might also depend on the probability of working at any other firm j', which will depend on the connections to j'.

Second, separately estimating employment and wage regressions cannot identify the causal effect of social connections on wages since it ignores selection. Because connections also impact the identity of the firms the workers end up working at (and for which we observe the wage information), there is a need to estimate the matching and wage-setting questions jointly.

Three, counterfactual policies might lead to equilibrium effects that cannot be captured by reduced-form estimation. For example, generating new connections between a set of workers and a set of firms might affect the structure of wages in the economy, which in turn will change the equilibrium matching.

The model addresses these concerns by 1) taking into account the full structure of connections in the economy, 2) jointly study the matching and wage-setting questions, and 3) doing it in an equilibrium framework.

During the model estimation and the counterfactual analysis, I rely on the same identification strategy I used above, comparing active and phantom connections. In the absence of an identification strategy, the estimated additional probability of knowing about a job or working in a firm given a meeting could also reflect the effects of job characteristics correlated with connections, such as location or industry. My identification strategy allows me to isolate the causal effect of connections in each mechanism. Likewise, due to the identification strategy, the counterfactuals identify what would happen if someone had more connections, not other factors associated with connections.

[35]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 opens. Solow et al. (2011) find no evidence that the Rooney Rule has increased the number of minority head coaches. This is not a surprise because search frictions are not likely to be significant in such a small market where everyone knows everyone else. A similar law in a large labor market will not necessarily produce the same result.

4.1 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 above, namely phantom, weak and strong connections. $c = 0$ denotes no connections.

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:

$$m_{ij} = 1\left(\rho_{ij} \leq p_{ij}\right) \tag{4}$$

where ρ_{ij} is a draw from an iid 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:

$$V_{ij} = f_{ij} - w_{ij} \tag{5}$$

where f_{ij} is the firm's surplus from the match (in terms of dollars), and w_{ij} is the wage the firm pays to the worker. The utility of workers is simply the wage they get:

$$U_{ij} = w_{ij}. \tag{6}$$

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 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. In particular, the model relaxes the "separability assumption" which is in use in the majority of this literature (Salanié 2015; Chiappori et al. 2017; Galichon and Salanié 2020).[36]

4.2 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:

$$\sum_{j=1}^{J} \mu(i,j) \leq 1 \quad , \quad \sum_{i=1}^{I} \mu(i,j) \leq 1 \quad , \quad \mu(i,j) = 1 \implies m(i,j) = 1 \tag{7}$$

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

$$\mu(i,j) = 1 \implies j \in \text{argmax}_{j \in m_i} U_{ij} \quad \text{and} \quad i \in \text{argmax}_{i \in m_j} V_{ij} \tag{8}$$

In words, a pairwise stable match in this context means that no pair of unmatched workers and firms who have previously met strictly prefer each other.

4.2.1 Finding the equilibrium matching

Let $f_{ij} = U_{ij} + V_{ij}$ be the joint surplus from a match between worker i and firm j. Shapley and Shubik (1972) show that μ is an equilibrium matching if and only if it maximizes the total joint surplus

$$\begin{aligned} &\mu \in \text{argmax}_{\mu} \sum_{\mu(i,j)=1} f_{ij} \\ \text{s.t.} \quad &\mu \text{ is feasible, i.e., equation (7) holds} \end{aligned} \tag{9}$$

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 given the meeting constraints.[37] The assignment problem can be solved by

[36] See Fox et al. (2018) for a notable exception.

[37] The equilibrium matching is generically not efficient. A matching with a higher total surplus might involve a match between an employee and a firm that had not met earlier. However, the equilibrium matching is constrained efficient, given the meeting restriction.

linear programming or the auction algorithm. In practice, I find the auction algorithm much faster for the problem at hand[38].

4.2.2 Finding the equilibrium wages

Generally, if there exists a feasible matching, there exists a unique equilibrium matching (Shapley and Shubik 1972).[39] However, the equilibrium wages that support the equilibrium matching are not unique. First, if w is an equilibrium wage schedule, so is $w+r$.[40] Therefore, one needs to normalize the location of wages in each market.[41]

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

In words, the set of equilibrium wages is characterized by component-wise upper- and lower-bound wages. The upper bound wages correspond to the workers' preferred equilibrium, while the lower bound wages correspond to the firms' preferred equilibrium (Bonnet et al. 2018).[42]

Given the equilibrium matching, the bounds on the equilibrium wages can be found using the Bellman-Ford algorithm (see Appendix E.3). To get a unique prediction of the equilibrium wages, I assume the wages are:

$$w_i = \lambda \underline{w}_i + (1-\lambda)\bar{w}_i \tag{10}$$

for some $\lambda \in [0,1]$. In the main estimation of the model, i assume $\lambda = 1/2$. However, in Appendix E.7, I check the sensitivity of the results to different values of λ.

[38] Appendix E.2 describes the Auction algorithm and discusses its properties in further detail.

[39] This is true under standard regularity conditions. For example, if the joint surplus f_{ij} is drawn from a continuous distribution, then with probability one, the equilibrium matching is unique.

[40] 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 wages location. See Dupuy and Galichon (2014) for the case that singles are observed.

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

[42] In a standard matching model, when every worker can work at any firm, the set of equilibrium wages 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 the set of feasible matches. In this case, the set of equilibrium wages shrinks to a singleton only when the number of meetings per worker goes infinity. In practice, I simulate the model with a small number of meetings per worker; therefore, the set of equilibrium wages has a non-trivial range.

4.3 Parametrization

I assume a flexible model in which the meeting and surplus 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:

$$p_{ij} = p_{txyc}. \tag{11}$$

Likewise, the surplus of a firm j is:

$$log(f_{ij}) = b + \beta_{txyc} + \sigma \cdot \xi_{ij} \tag{12}$$

where β_{txyc} is the systematic surplus and depends on the observable characteristics of i and j, and ξ_{ij} is drawn from an iid standard normal distribution, independent of the meeting error term ρ_{ij}.[43] σ is a parameter that needs to be estimated. In line with the standard assumption in the labor economics literature that assumes an additive error in the log wage equations, I assume a log-linear specification of the systematic and idiosyncratic parts of the firm's surplus, which is closely related to the wages.

The meeting probability and the firm's systematic surplus 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 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.[44] Finally, similarly to the reduced-form estimation, I use 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.

Note that I use the AKM firm premiums only to classify firms. The model's "pay premium" of each bin of firms is estimated within the model and not based on the premiums

[43] Note, however, that the systematic parameters p_{txyc} and β_{txyc} can be correlated.

[44] In this part of the paper, I do not use the geographic location of the workers and firms for classification for computational reasons. However, differences in geographic location and other differences in observed and unobserved characteristics of the workers and the firms are netted out by focusing on the difference between real and phantom connections.

estimated in the AKM model. Likewise, I do not rely on the AKM-style log-additive assumption in worker's and firm's effects anywhere in the estimation but estimate a separate surplus parameter for each $txyc$ cell.

4.4 Moments

There are three sets of parameters in the model that need to be estimated: the firm's systematic surplus β_{txyc}, the meeting probability 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 previous sections. 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.

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 calculated. Using the equilibrium outcome, I can compute the model analogs to the data moments $\hat{h}(\theta; \zeta) = (\hat{\mu}_{txyc}(\theta; \zeta), \hat{w}_{txyc}(\theta; \zeta), \hat{Var}_w(\theta; \zeta))$.[45]

4.5 Estimation

The large number of parameters in the model does not allow estimation using indirect search methods such as the method of simulated moments. I use an update mapping to "invert" the observed matches and wages into the parameters to estimate the model. In each iteration, the algorithm updates the parameters based on comparing 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

[45]See Appendix E.4 for discussion about the identification of the model.

mapping:[46]

$$\beta_{txyc}^{h+1} = \beta_{txyc}^{h} + \eta \left[log(\mu_{txyc} \cdot w_{txyc}) - log(\hat{\mu}_{txyc}(p^h, \beta^h, \sigma^h, b^h) \cdot \hat{w}_{txyc}(p^h, \beta^h, \sigma^h, b^h)) \right] \quad (13)$$
$$p_{txyc}^{h+1} = p_{txyc}^{h} + \eta \left[log(\mu_{txyc}) - log(\hat{\mu}_{txyc}(p^h, \beta^h, \sigma^h, b^h)) \right] \quad (14)$$
$$\sigma^{h+1} = \sigma^h + \eta \left[log(WithinVar_w) - log(Within\hat{V}ar_w(p^h, \beta^h, \sigma^h, b^h)) \right] \quad (15)$$
$$b^{h+1} = b^h + \eta \left[log(Var_w) - log(V\hat{a}r_w(p^h, \beta^h, \sigma^h, b^h)) \right] \quad (16)$$

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. Finally, β_{txyc}^h, p_{txyc}^h, σ^h, and b^h are the parameters in iteration h.

The update equations are defined using insights about the relationships between parameters and moments, which are discussed in detail in Appendix E.4. Starting with the match surplus parameter in equation (13), a higher surplus 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 positively impacts the share of matches but does not significantly impact the average wage within a cell. Hence, it is updated only by the share of matches (equation (14)).

Two additional parameters that need to be estimated are the idiosyncratic surplus parameter σ and the surplus constant b, which are updated by the within-group wage variance $WithinWageVar$ and overall variance $WageVar$ (equations (15)-(16)). I add the surplus 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 surplus 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 surplus 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 surplus and σ in the right direction.

In sum, this section suggests a novel estimation procedure to estimate two sets of unobserved model 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 estimation procedure extends the contraction mapping algorithm proposed by Berry et al. (1995) to "invert" one

[46]For computational reasons explained below, I add the surplus constant b explicitly to the estimation process and normalize the mean of β_{txyc} (weighted by μ_{txyc}) to zero. Likewise, I explicitly add the within-group wage variance to the set of moments (besides the overall wage variance). 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. Further details about the estimation can be found in Appendix E.5.

set of data moments (market shares) into one set of parameters (utilities). This directed updating procedure enables estimating models with many parameters, even when the simulation of each model's iteration is expensive, in cases where the theory can guide us about the relationships between the parameters and the moments.

5 Model results

I estimate the model 100 times with different values of the shocks ζ and present the average results (and their standard errors) across the model's 100 sets of estimated parameters.[47]

5.1 Impact of parental connections

To summarize the model estimates, I project the estimated parameters on workers', firms', and connections' characteristics. Table 4 reports the WLS estimates of the equation:

$$\theta_{txyc} = b + \delta_c + \gamma_1 Arab_x + \gamma_2 Female_x + \gamma_3 College_x + \psi_y + \epsilon_{txyc} \tag{17}$$

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 surplus 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.

First, I study the contribution of the characteristics of connections, workers, and firms to the meeting parameters by estimating equation (17) with $log(p_{txyc})$ as the outcome. The effect of all types of connections on meeting probability is positive and significant (Table 4, column 1). 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 2.1 and 5.9, respectively.

Next, I estimate equation (17) with β_{txyc} as an outcome. The second column of Table 4 shows that phantom connections only slightly affect the surplus parameter (1.2 log points, not statistically significant). Weak and strong connections increase the estimated surplus 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

[47]The model fit and precision are very good. See Appendix E.6 for details.

the surplus parameter by 2.8 and 14.6 log points, respectively.

The causal impact of weak connections on match surplus is translated into an increase of 35% in the likelihood of a match given a meeting.[48] The impact of this channel on matching is smaller in magnitude compared to the effect of the first channel (115% increase). Combining the two effects implies that workers are 2.9 times more likely to find employment in firms with weak parental connections than phantom-connected firms. This effect is somewhat smaller than the reduced form estimate (odds ratio of 3.7).

The differences in the firm surplus from connected and not connected hiring should not 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 surplus from hiring a connected worker might be higher because of a lower uncertainty about the productivity of the worker or the match. This lower uncertainty, 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 that exists in the data between connections and tenure at the first job.[49]

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 surplus is monotonically increasing with the job type, as expected.

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 A9 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

[48]I obtain this result using simulations comparing the probability of working in a firm with and without the match surplus associated with connections. See Section 6.2 for further details.

[49]Richer data are needed to estimate two or more of these sub-channels separately. 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.

is stronger for Arabs than for Jews.[50]

6 Counterfactuals

6.1 Causal connections

To get the causal effect of connections (net of the impact of confounders), I exploit the identification strategy from the previous 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 surplus of a weak "causal" connection for workers of type x, firms of type y, and year t is:

$$\beta_{txy,weak}^{causal} = \beta_{txy,none} + \beta_{txy,weak} - \beta_{txy,phantom}. \tag{18}$$

where $\beta_{txy,c}$ is the estimated systematic surplus of that txy group with connections of type $c \in \{none, phantom, weak\}$.To put it another way, I measure the excess effect of connections on the surplus net of confounders correlated with connections by the difference between the estimates with weak and phantom connections. Likewise, the meeting probability of a weak "causal" connection is:

$$p_{txy,weak}^{causal} = p_{txy,none} \cdot p_{txy,weak}/p_{txy,phantom} \tag{19}$$

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.

6.2 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 surplus of the affected workers with and without the additional connection/meeting. The surplus difference measures the wage-equivalence 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. First, I add a new meeting between a random worker and a random firm assuming the systematic surplus associated with unconnected pairs. Second, I add the surplus associated with weak causal connections to an existing meeting. This

[50] The results are robust for different values of the bargaining parameter λ (Appendix E.7).

exercise isolates the effect of the surplus channel alone. Finally, I add a new meeting with the assumption that the worker and firm have a weak causal connection.

The first column of Table 5 reports the results of this exercise with 100,000 new meetings/connections (1,000 for each of the 100 sets of estimated parameters of the model). For convenience, I report all results in terms of percentages of new workers' average wage. The average value of one additional meeting without the surplus effect is 2.2 percent of new workers' average wage. Adding connections to a random existing meeting, the wage increases by 1.5 percent. Finally, adding a new meeting is with a causal weak connected firm increases the wage by 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 surplus 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 percent of the average wage (Table 5, row 1).

If we add the surplus 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 of the average wage . 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 5, row 2).

Finally, if we assume that the new meeting is accompanied by the surplus of a causal weak connection, the probability that the workers will work 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 6.6 percent of the cases, the wage changes without a job change. These events yield average gains of 9 percent of the average wage (Table 5, 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.

Using the simulation results, I can also translate the impact of connections on match surplus into matching probabilities. Specifically, given a meeting, the likelihood of working in a random firm without the surplus effect of connections is 4.0 percent. However, the probability of working at the same firm with the surplus effect of connections is 5.5 percent. Taken together, having a causal weak connection at a firm increases the probability of a match by 35 percent.

Not all meetings/connections are equal. Figure A11 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 surplus effect, an existing meeting with the surplus effect, and a new meeting with the surplus effect).

6.3 Between-group pay gaps

Social connections might not only be important for individuals, but also for the society at large, in particular for income inequality (Calvo-Armengol and Jackson 2004; Bolte et al. 2020). In what follows, I use the structural model to examine how much of the pay gap between different groups in Israel is due to differences in the quality and quantity of connections people inherit from their parents. I do it in two ways. First, I check the predicted inequality if the different groups, Arabs and Jews or males and females, would have similar quantities and qualities of connections. Second, 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, the same gender and education characteristics, and 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 equals.

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 counterfactual exercise naively assumes that new connections' meeting and surplus parameters are the estimated parameters of real connections (either weak or strong) of the corresponding $txyc$ cells. By that, it ignores the fact that these estimates combine the causal impact of connections with confounders. However, the counterfactual exercise with the identification strategy correctly assumes that the new connections have only the excess effect of real connections relative to phantom connections, as defined above (equation (18) and (19), and the analogous definition for strong connections).

Starting with the ethnic pay gap, the first row of Table 6 shows the results when the share of connections with all firms is equal for Arabs and Jews. The benchmark gap in wage between Arabs and Jews is 502 NIS or 8.4 percent of the average wage. Without the identification strategy, the estimated reduction in the ethnicity pay gap is 59.5, 0.4, and 67.6 percent, given the meeting effect, surplus effect, and both effects, respectively.

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 5.1, 1.1, and 11.7 percent, given the meeting effect, surplus effect, and both effects, respectively. The large difference between the counterfactual results with and without the identification strategy indicates the importance of using "causal" variation in structural estimation and interference. Without the identification strategy, we wrongly attribute the impact of confounders, 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.

The results of these counterfactual exercises are informative about the effectiveness of different policies in reducing inequality. For example, consider a policy that increases the number of job interviews of Arab candidates for open positions at some firms. This policy is equivalent to increasing the number of connections between the candidates and the firms but only considering the impact of connections on the meeting rates. Suppose this policy is tuned such that the minimum job-interview requirements of Arab candidates exactly replace the missing (causal) connections of Arabs compared to Jews. In that case, the wage gap will decrease by 5.1 percent, according to the model. However, other policies, such as subsidizing internships between Arabs candidates and firms, might also impact the match values. In that case, the ethnic pay gaps would decrease by as much as 11.7 percent.

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 increases by 2.3 percent. However, using the identification strategy, the gap increases by 0.1 percent, and the change is not statistically significantly different from zero (Table 6, 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, strong connections only, and strong and weak connections together. Panel B of Table 6 shows that prohibiting the hiring of workers with connections, as some anti-nepotism rules do, *increases* the predicted ethnic pay gap by 8.9 percent if only weak connections are prohibited, by 44.3 percent if only strong connections are prohibited, and by 56.4 percent if both weak and strong connections are prohibited. The gender pay gap *declines* by 4.0, 20.3, and 25.3 percent, respectively, in these different scenarios.

The difference between the results of the two scenarios can be explained by considering the differences in the quality of connections and 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, the higher impact of connections, obtained both in the reduced-form and structural estimation, indicate that Arabs rely more heavily on connections compared to Jews (Figures 3 and A9). 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. 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.

7 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 of the new workers 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.

Estimates of the structural model show 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 explain a large proportion of Israel's ethnic pay gap. Equalizing the quantity and quality of Arabs' and Jews' connections decreases the ethnic pay gap 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.

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 with fewer connections or policies requiring interviews of these candidates for open positions. 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, policies aim to increase the number of job interviews 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 *increase* inequality, as workers from disadvantaged 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. First, like most of the matching literature, the model is static. Estimating a dynamic version of the model will enable studying 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 over time allows estimating workers' and firms' fixed effects, which cannot be separately identified in a static model. Second, having information on other labor market outcomes could allow the estimation of additional unobserved parameters, such as the workers' non-wage match surplus 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 essential in crafting policies to help reduce inequity in the labor market.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

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

Ahyja, R. K., J. B. Orlin, and T. L. Magnanti (1993). *Network flows: theory, algorithms, and applications*. Prentice-Hall.

Arellano-Bover, J. (2020). Career consequences of firm heterogeneity for young workers: First job and firm size. *Unpublished*.

Arellano-Bover, J. and F. Saltiel (2021). Differences in on-the-job learning across firms. *Unpublished*.

Arellano-Bover, J. and S. San (2020). The Role of Firms in the Assimilation of Immigrants. *Unpublished*.

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

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

Banfi, S. and B. Villena-Roldan (2019). Do high-wage jobs attract more applicants? Directed search evidence from the online labor market. *Journal of Labor Economics 37*(3), 715–746.

Bayer, P., S. L. Ross, and G. Topa (2008). Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of Political Economy 116*(6), 1150–1196.

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

Becker, G. S. (1973). A theory of marriage: Part I. *Journal of Political economy 81*(4), 813–846. Publisher: The University of Chicago Press.

Berman, E. (2000). Sect, subsidy, and sacrifice: an economist's view of ultra-orthodox Jews. *The Quarterly Journal of Economics 115*(3), 905–953.

Bernard, F., N. Vlassis, P. Gemmar, A. Husch, J. Thunberg, J. Goncalves, and F. Hertel (2016). Fast correspondences for statistical shape models of brain structures. In *Medical Imaging 2016: Image Processing*, Volume 9784, pp. 97840R. International Society for Optics and Photonics.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–890.

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

Bolte, L., N. Immorlica, and M. O. Jackson (2020). The Role of Referrals in Inequality, Immobility, and Inefficiency in Labor Markets. *Unpublished*.

Bonhomme, S., T. Lamadon, and E. Manresa (2019). A distributional framework for matched employer employee data. *Econometrica 87*(3), 699–739.

Bonnet, O., A. Galichon, K. O'Hara, and M. Shum (2018). Yogurts Choose Consumers? Estimation of Random Utility Models via Two-Sided Matching. *Unpublished*.

Burdett, K. and D. T. Mortensen (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, 257–273.

Caldwell, S. and O. Danieli (2021). Outside options in the labor market. *Unpublished*.

Caldwell, S. and N. Harmon (2019). Outside Options, Bargaining, and Wages: Evidence from Coworker Networks. *Unpublished*.

Calvo-Armengol, A. and M. O. Jackson (2004). The effects of social networks on employment and inequality. *American Economic Review 94*(3), 426–454.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics 36*(S1), S13–S70.

CBS (2019). Statistical abstract of Israel. Technical report, Central Bureau of Statistics, Jerusalem, Israel.

Chiappori, P.-A. and B. Salanié (2016). The econometrics of matching models. *Journal of Economic Literature 54*(3), 832–61.

Chiappori, P.-A., B. Salanié, and Y. Weiss (2017). Partner choice, investment in children, and the marital college premium. *American Economic Review 107*(8), 2109–67.

Choo, E. and A. Siow (2006). Who marries whom and why. *Journal of Political Economy 114*(1), 175–201.

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

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

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

Del Boca, D., C. Flinn, and M. Wiswall (2014). Household choices and child development. *Review of Economic Studies 81*(1), 137–185.

Demange, G. and D. Gale (1985). The strategy structure of two-sided matching markets. *Econometrica 53*(4), 873–888.

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

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

Dupuy, A. and A. Galichon (2015). A Note on the Estimation of Job Amenities and Labor Productivity. *Unpublished*.

Dustmann, C., A. Glitz, U. Schönberg, and H. Brücker (2016). Referral-based job search networks. *The Review of Economic Studies 83*(2), 514–546.

Eliason, M., L. Hensvik, F. Kramarz, and O. N. Skans (2022). Social connections and the sorting of workers to firms. *Journal of Econometrics*. Publisher: Elsevier.

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

Fox, J. T., C. Yang, and D. H. Hsu (2018). Unobserved heterogeneity in matching games. *Journal of Political Economy 126*(4), 1339–1373.

Fuchs, H. and G. Epstein (2019). The Labor Market: An Overview. *State of the Nation: Policy, Economy and Society in Israel*, 139–162.

Galichon, A. and B. Salanié (2020). Cupid's invisible hand: Social surplus and identification in matching models. *Unpublished*.

Galor, O. and D. Tsiddon (1997). Technological progress, mobility, and economic growth. *American Economic Review 87*(3), 363–382.

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

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

Hagedorn, M., T. H. Law, and I. Manovskii (2017, January). Identifying Equilibrium Models of Labor Market Sorting. *Econometrica 85*(1), 29–65.

Hensvik, L. and O. N. Skans (2016). Social networks, employee selection, and labor market outcomes. *Journal of Labor Economics 34*(4), 825–867.

Jaffe, S. and S. Weber (2019). The effect of meeting rates on matching outcomes. *Economic Theory 67*(2), 363–378.

Jager, S. (2016). How substitutable are workers? evidence from worker deaths. *Unpublished*.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of political economy 87*(5, Part 1), 972–990. Publisher: The University of Chicago Press.

Kramarz, F. and O. N. Skans (2014). When strong ties are strong: Networks and youth labour market entry. *Review of Economic Studies 81*(3), 1164–1200.

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

McPherson, M., L. Smith-Lovin, and J. M. Cook (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology 27*(1), 415–444.

Menzio, G. and S. Shi (2011). Efficient search on the job and the business cycle. *Journal of Political Economy 119*(3), 468–510. Publisher: University of Chicago Press Chicago, IL.

Miller, C. and I. M. Schmutte (2021). The Dynamics of Referral Hiring and Racial Inequality: Evidence from Brazil. Technical report, National Bureau of Economic Research.

Montgomery, J. D. (1991). Social networks and labor-market outcomes: Toward an economic analysis. *American Economic Review 81*(5), 1408–1418.

Mortensen, D. T. and C. A. Pissarides (1994). Job creation and job destruction in the theory of unemployment. *The Review of Economic Studies 61*(3), 397–415.

Noy, C. and E. Cohen (2005). *Israeli backpackers: From tourism to rite of passage.* SUNY Press.

Oreopoulos, P., T. Von Wachter, and A. Heisz (2012). The short-and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics 4*(1), 1–29.

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

Postel–Vinay, F. and J.-M. Robin (2002). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica 70*(6), 2295–2350.

Rees, A. (1966). Information networks in labor markets. *American Economic Review 56*(1/2), 559–566.

Rotemberg, J. J. (1994). Human relations in the workplace. *Journal of Political Economy 102*(4), 684–717. Publisher: The University of Chicago Press.

Roussille, N. (2020). The central role of the ask gap in gender pay inequality. *Unpublished*.

Salanié, B. (2015). Identification in Separable Matching with Observed Transfers. *Unpublished*.

Shapley, L. S. and M. Shubik (1972). The assignment game I: The core. *International Journal of Game Theory 1*(1), 111–130.

Shi, S. (2002). A directed search model of inequality with heterogeneous skills and skill-biased technology. *The Review of Economic Studies 69*(2), 467–491. Publisher: Wiley-Blackwell.

Shimer, R. and L. Smith (2000). Assortative matching and search. *Econometrica 68*(2), 343–369.

Solow, B. L., J. L. Solow, and T. B. Walker (2011). Moving on up: The Rooney rule and minority hiring in the NFL. *Labour Economics 18*(3), 332–337.

Staiger, M. (2021). The Intergenerational Transmission of Employers and the Earnings of Young Workers. *Unpublished*.

Stinson, M. and C. Wignall (2018). Fathers, Children, and the Intergenerational Transmission of Employers. *Unpublished*.

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	0.007
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Weak connections	0.051	0.031	0.144	0.057	0.032
	(0.002)	(0.002)	(0.007)	(0.002)	(0.004)
Strong connections	0.490	0.368	0.922	0.505	0.445
	(0.008)	(0.008)	(0.023)	(0.010)	(0.016)
R0 (no connections)	0.005	0.005	0.006	0.005	0.006
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ratio weak-phantom	3.706	3.307	4.188	3.957	2.971
	(0.214)	(0.247)	(0.319)	(0.247)	(0.371)
Ratio strong-phantom	32.85	34.51	25.93	32.36	35.41
	(1.467)	(2.010)	(1.848)	(1.694)	(3.571)
Observations	20,936,981	16,654,016	4,282,965	15,162,471	5,774,510
N firms	148,066	142,545	116,514	144,302	133,004
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	40,466,632	32,976,991	7,489,641	31,412,673	9,053,959

Notes: This table reports the mean (and standard deviation) of the estimated coefficients of phantom, weak, and strong connections across 100 estimations of equation (1) using a 20 percent random sample of workers each time. The employment outcome is scaled by 100. 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 sub-groups of workers.

Table 3: Correlation between parental connections at first job and salary, job tenure, and wage growth

	Log salary		Job tenure		Salary growth	
	(1)	(2)	(3)	(4)	(5)	(6)
Phantom connections	-0.007	0.012	0.123	0.098	-0.006	0.008
	(0.005)	(0.004)	(0.021)	(0.021)	(0.010)	(0.010)
Weak connections	0.018	0.026	0.182	0.187	-0.052	-0.020
	(0.005)	(0.004)	(0.023)	(0.022)	(0.011)	(0.012)
Strong connections	0.074	0.083	0.601	0.441	-0.014	0.012
	(0.004)	(0.003)	(0.024)	(0.018)	(0.005)	(0.006)
Diff. weak-phantom	0.025	0.014	0.058	0.089	-0.045	-0.027
	(0.006)	(0.005)	(0.031)	(0.030)	(0.014)	(0.016)
Diff. strong-phantom	0.081	0.071	0.477	0.343	-0.007	0.004
	(0.005)	(0.004)	(0.029)	(0.027)	(0.011)	(0.012)
Mean (no connections)	8.577	8.577	1.933	1.933	0.313	0.313
Fixed effects	Group	Group + Firm	Group	Group + Firm	Group	Group + Firm
Observations	220,806	220,806	220,806	220,806	106,368	106,368
N firms	54,321	54,321	54,321	54,321	35,335	35,335
R^2 (full model)	0.169	0.624	0.127	0.414	0.060	0.427
R^2 (projected model)	0.004	0.006	0.014	0.007	0.000	0.000

Notes: This table reports the correlation between parental connections at the first job on the one hand and salary and tenure at the first job, and salary growth in the first three years, on the other hand. The outcome in columns 1-2 is (log) monthly salary in the first year of the first job. The outcome in columns 3-4 is the number of sequential years workers worked at their first job (truncated in 2015). The outcome in columns 5-6 is the wage growth after three years in the labor market (only for the 2006-2012 cohorts). Columns 1,3 and 5 include group fixed effects. Columns 2,4 and 6 include group and firm fixed effects. Groups are constructed using all combinations of the workers' observable characteristics (ethnicity, education, gender, year of the first job, age, and district of residence). "Diff. weak-phantom" is the difference between the coefficients of weak and phantom connections. "Diff. strong-phantom" is defined similarly. "Mean (no connections)" is the mean of the dependent variable for workers with no connections at their first job. Robust standard errors clustered by the firm are reported in parentheses.

Table 4: Projection of the model estimates on workers', firms', and connections' characteristics

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

Notes: This table reports the results of regressing the meeting and surplus estimates 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 5: 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 surplus 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 surplus 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 surplus 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 (in terms of percentages of new workers' average wage). Each row reports the average change in the salary of workers in one of three 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 surplus to reflect the surplus of a causal weak connection, and 3) adding a random meeting with causal weak connections. The surplus of a causal weak connection is the excess surplus 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, based on 1,000 new meetings/connections for each estimation, and the table reports the averages across the 100 estimations (and their standard errors in parentheses).

Table 6: Counterfactual impacts of connections on between-group pay gaps

A. Equalizing number of connections per worker

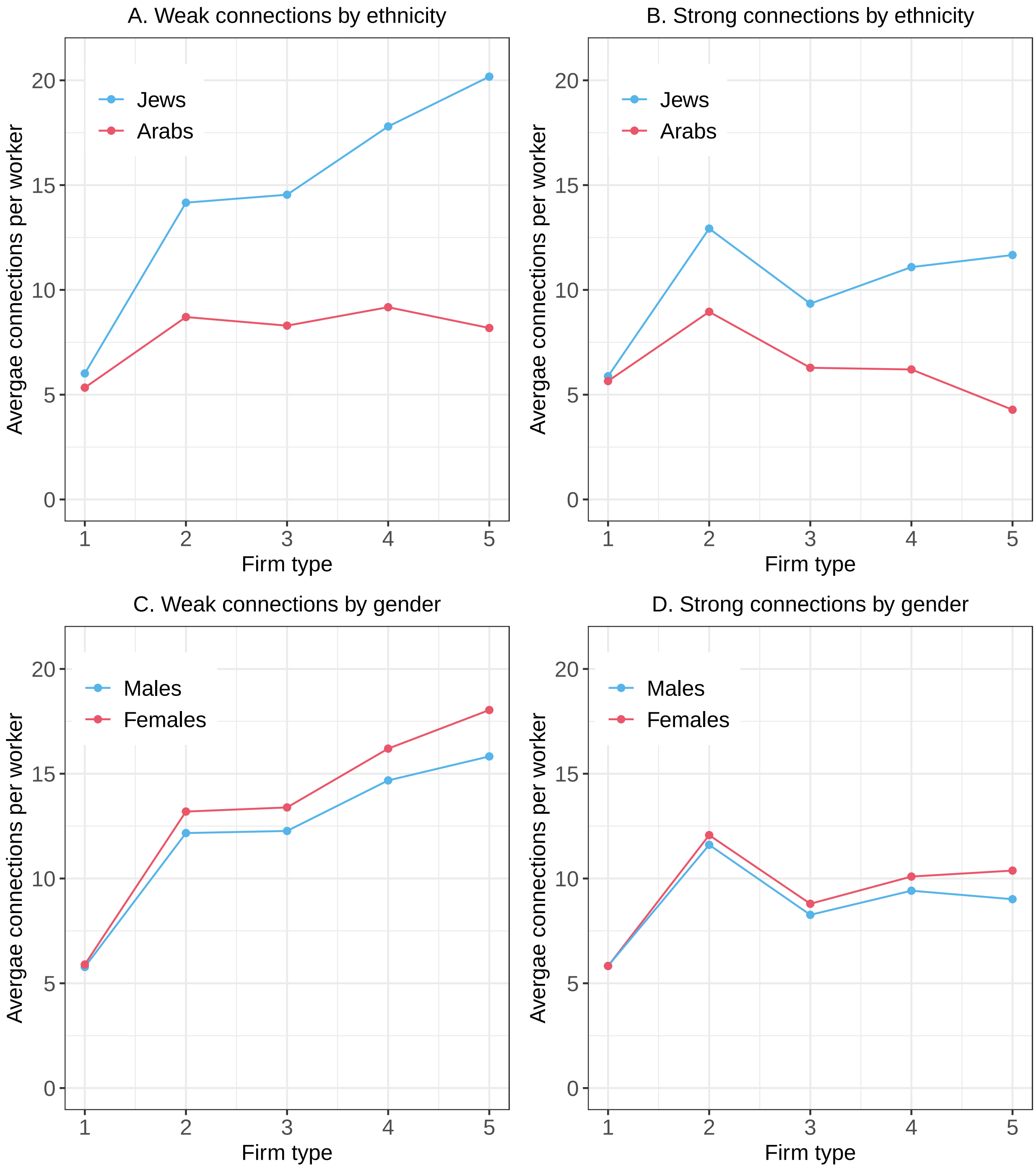
	Gap (% Average)	Without identification strategy			With identification strategy		
		Meetings effect	Surplus effect	Both effects	Meetings effect	Surplus effect	Both effects
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethnicity gap	-8.4	-59.5	-0.4	-67.6	-5.1	-1.1	-11.7
	(0.351)	(4.866)	(0.168)	(3.031)	(0.679)	(0.297)	(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 (% Average)	Weak	Strong	Weak + strong
	(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)

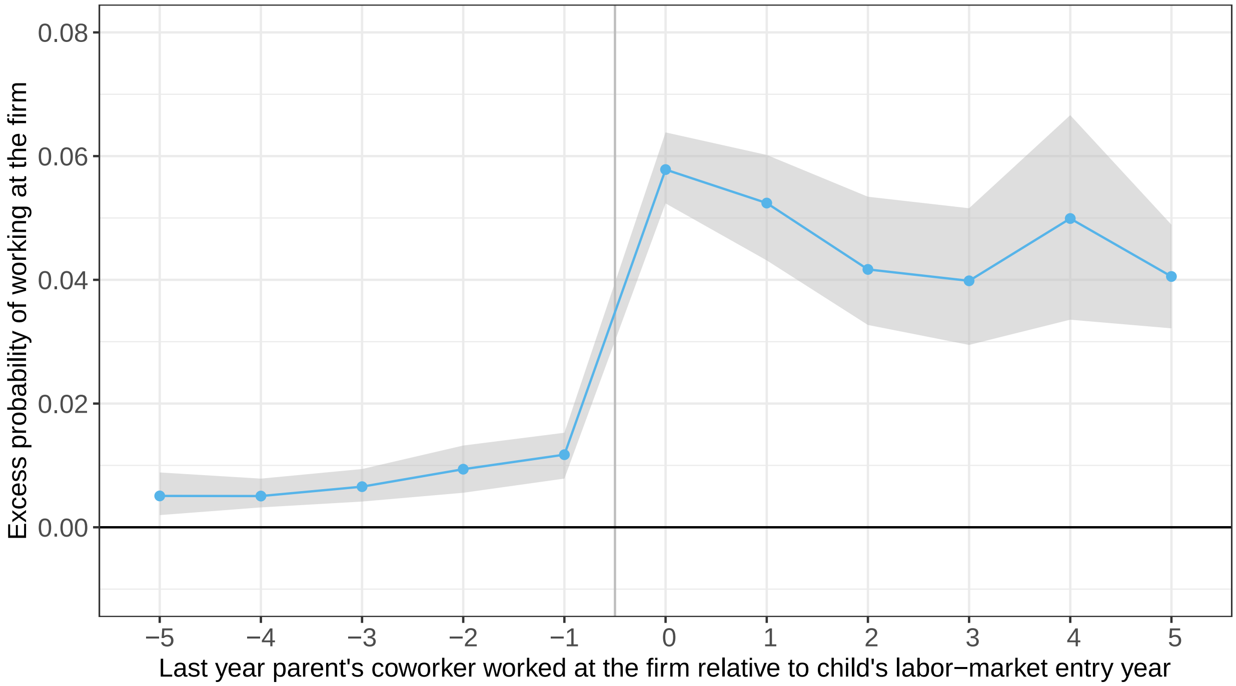
Notes: This table shows the contribution of parental connections to the ethnic and gender pay gaps in two scenarios. Panel A reports estimates from equalizing the connections between the ethnic and gender group. Specifically, in the first row, I present the ethnic pay gap predicted by the model assuming 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 connections (either weak or strong) have the same impact on the meeting rate and the match surplus as a real connection of the same type in the same *txyc* cell. In columns (6)-(8), I assume that the impact of new connections on the meeting rate and the match surplus is the excess impact of strong or weak connections on these parameters compared to phantom connections ("causal connections"). In columns (2) and (5), I shut down the surplus effect of new connections (assuming they are similar to the surplus 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



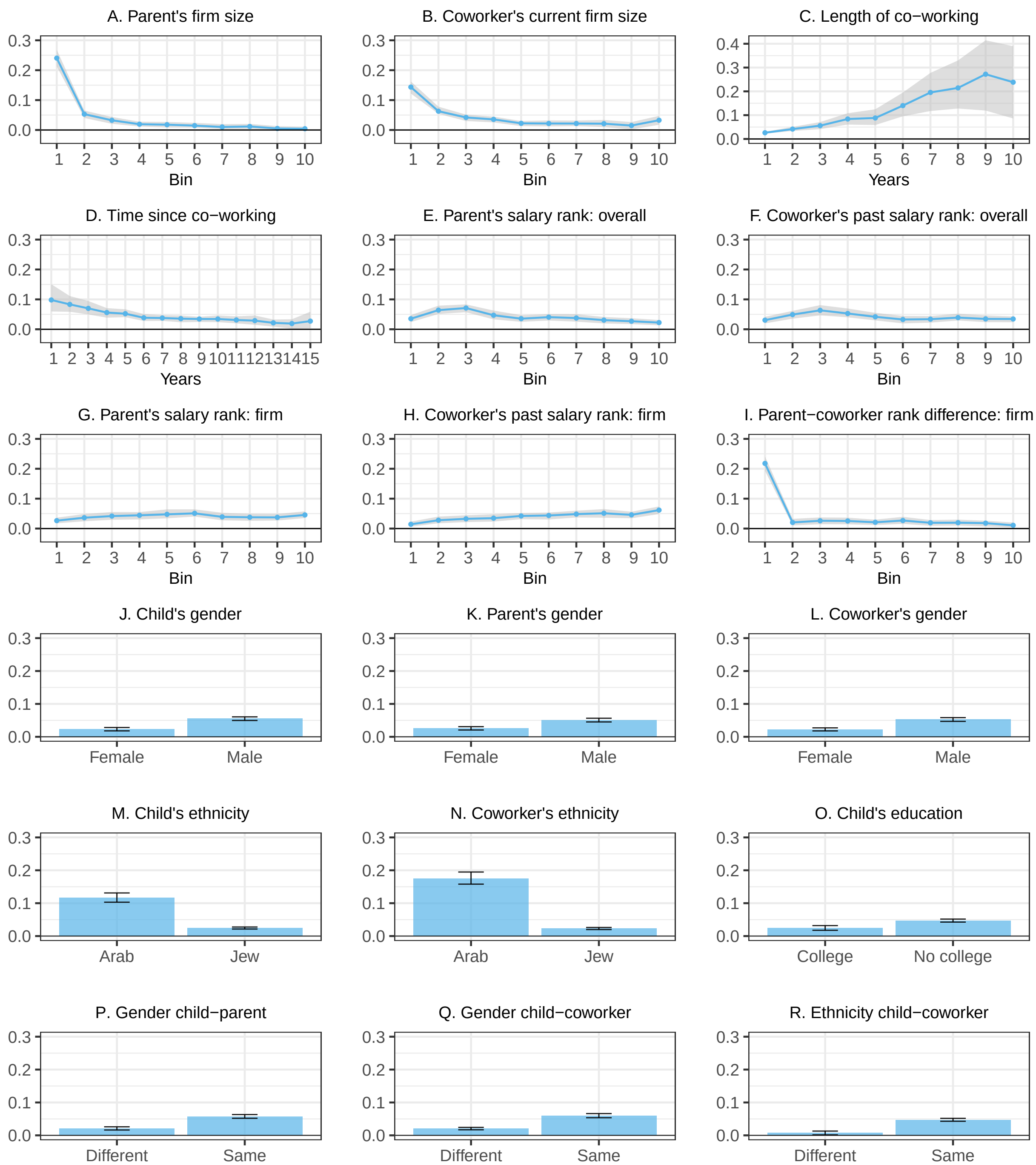
Notes: This figure shows the average number of weakly and strongly connected firms per worker by workers' ethnicity and gender, and by quintiles of the AKM firm premium, averaged over the years 2006-2015.

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 figure shows the mean (and 95 percent confidence interval) of the estimated coefficients of phantom and weak connections across 100 estimations of equation (2) using a 20 percent random sample of workers each time. The employment outcome is scaled by 100. The vertical line between -1 and 0 indicates the change from worker-firm pairs with phantom connections to pairs with weak connections.

Figure 3: 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 phantom-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 and phantom connections, using a 20 percent random sample of workers each time. The employment outcome is scaled by 100. I construct the bounds of the 95 percent confidence intervals using the 2.5 and 97.5 percentiles of that distribution of coefficients.

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: Ethnicity and gender pay gaps: workers at ages 22-69, 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 22-69 in 2015. The outcome variable is the log of the average monthly wage in 2015. All columns include two dummy variables indicate if the worker is Arab or female, respectively. All columns also include a set of dummy variables for every combination of age, education, and the residential district in 2015. Columns 2 also includes a full set of firm fixed effects. Robust standard errors clustered by group (age-education-district) and firm are reported in parentheses.

Table A3: 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 two dummy variables indicate if the worker is Arab or female, 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-education-age-district) and firm are reported in parentheses.

Table A4: 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	
	(0.001)	(0.001)	
Phantom (single + multiple contacts)			0.015
			(0.001)
Weak (signle contact)	0.051	0.053	
	(0.002)	(0.002)	
Weak (single + multiple contacts)			0.095
			(0.002)
Strong (direct + multiple contacts)	0.490		
	(0.008)		
Direct		3.091	3.092
		(0.059)	(0.059)
Multiple contacts		0.171	
		(0.005)	
R0 (no connections)	0.005	0.005	0.005
	(0.000)	(0.000)	(0.000)
Observations (firms x groups)	20,936,981	21,166,443	21,166,443
N firms	148,066	149,729	149,729
N groups	2,959	2,959	2,959
N workers	220,684	220,684	220,684
N connections	40,466,632	40,827,833	40,827,833

Notes: This table checks the robustness of the baseline results to alternative definitions of parental connections. table reports the mean (and standard deviation) of the estimated coefficients of parental connections across 100 estimations of equation (1) with separate coefficient for each type of parental connection using a 20 percent random sample of workers each time. The employment outcome is scaled by 100. 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 A5: Event-study plot of coefficients: Effect of parental connections on firm assignment

	Employment		
Phantom connections		Weak connections	
-5	0.005	0	0.058
	(0.002)		(0.003)
-4	0.005	1	0.052
	(0.001)		(0.004)
-3	0.007	2	0.042
	(0.001)		(0.006)
-2	0.009	3	0.040
	(0.002)		(0.006)
-1	0.012	4	0.050
	(0.002)		(0.009)
1	0.026	5	0.041
	(0.003)		(0.005)
2	0.017	Strong connections	
	(0.002)		
3	0.013		0.490
	(0.002)		(0.008)
4	0.009		
	(0.002)		
5	0.009		
	(0.002)		

Notes: This table reports the mean (and standard deviation) of the estimated coefficients of parental connections across 100 estimations of equation (2) using a 20 percent random sample of workers each time.

Table A6: 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
	(0.004)	(0.000)
Weak connections	-0.368	0.076
	(0.003)	(0.001)
Strong connections	-0.926	0.281
	(0.009)	(0.001)
R0 (no connections)	10.102	0.033
	(0.007)	(0.000)
Ratio weak-phantom	1.000	0.989
	(0.000)	(0.003)
Ratio strong-phantom	0.943	2.871
	(0.001)	(0.010)
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 compares the geographical distance between a worker and a firm and the probability that a firm belongs to the same 3-digit industry of the worker's parent for firms with parental connections relative to non-connected firms. The table reports the mean (and standard deviation) of the estimated coefficients of phantom, weak, and strong connections across 100 estimations of equation (1) with the outcome variables mentioned using a 20 percent random sample of workers each time. 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 A7: 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
	(0.016)	(0.010)	(0.008)
Phantom (Other)	0.010	0.010	0.010
	(0.001)	(0.001)	(0.001)
Weak (D/R)	0.065	0.032	0.041
	(0.033)	(0.016)	(0.015)
Weak (Other)	0.050	0.051	0.051
	(0.002)	(0.002)	(0.002)
Strong	0.487	0.487	0.487
	(0.008)	(0.008)	(0.008)
R0 (no connections)	0.005	0.005	0.005
	(0.000)	(0.000)	(0.000)
Ratio weak-phantom (D/R)	2.567	3.913	2.773
	(2.219)	(5.313)	(3.861)
Ratio weak-phantom (Other)	3.679	3.680	3.691
	(0.209)	(0.212)	(0.214)
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 mean (and standard deviation) of the estimated coefficients of phantom, weak, and strong connections across 100 estimations of equation (1) with separate coefficients for "D/R" connections ("death", "retirement" or both, depending on the column) and "other" phantom and weak connections, using a 20 percent random sample of workers each time. "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. R0 is the average probability of working in a non-connected firm. "Ratio weak-phantom: D/R" is the estimated odds ratio between working in a "D/R" weakly-connected firm and working in a "D/R" phantom-connected firm. "Ratio weak-phantom: Other" is defined similarly.

Table A8: 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
	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)
Weak connections	0.000	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.003)	(0.001)	(0.002)
Strong connections	0.000	0.000	0.001	0.000	0.000
	(0.003)	(0.002)	(0.010)	(0.004)	(0.005)
R0 (no connections)	0.007	0.006	0.011	0.008	0.007
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ratio weak-phantom	1.010	1.000	1.053	1.011	1.017
	(0.152)	(0.149)	(0.337)	(0.176)	(0.239)
Ratio strong-phantom	1.047	1.029	1.107	1.065	1.036
	(0.416)	(0.375)	(0.976)	(0.499)	(0.637)
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 mean (and standard deviation) of the estimated coefficients of phantom, weak, and strong connections across 100 estimations of equation (1) based on the new (randomized) data using a 20 percent random sample of workers each time. The employment outcome is scaled by 100. 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 sub-groups of workers.

Table A9: Moments-parameters elasticities

	Matches-surplus $dln(\mu)/d\beta$ (1)	Matches-meetings $dln(\mu)/dln(p)$ (2)	Wages-surplus $dln(w)/d\beta$ (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 surplus 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 $x' \neq x$. Each statistic is calculated separately for each of the 100 estimations of the model, based on 1,000 new meeting/surplus parameters for each estimation, and the table reports the averages across the 100 estimations (and their standard errors in parentheses).

Table A10: Model's fit, Model's precision, and Monte Carlo simulation

A. Model's fit

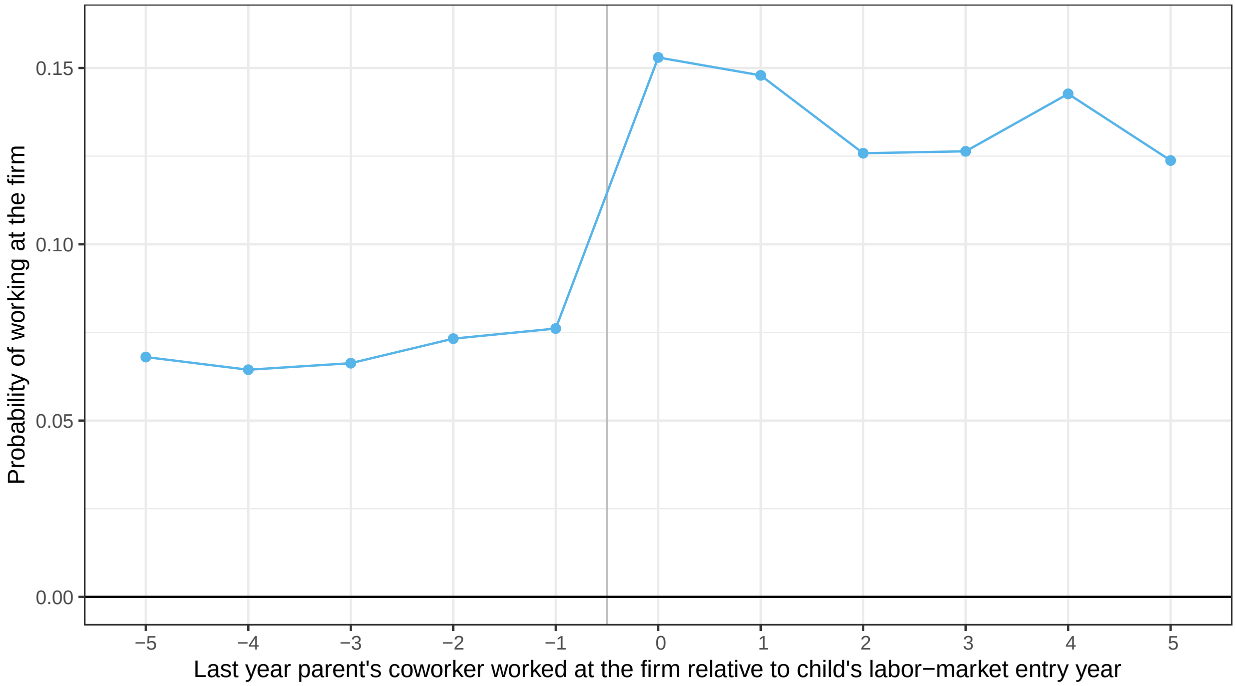
	Matches (μ_{txyc})	Av. wage (w_{txyc})	Overall wage variance	Within-group wage variance
	(1)	(2)	(3)	(4)
Abs. deviation	0.013	0.008	0.0008	0.0007
	(0.0006)	(0.0006)	(0.0006)	(0.0005)
Correlation	1.000	0.998		
	(0.00002)	(0.0002)		

B. Model's precision and Monte Carlo simulation

	Surplus (β_{txyc})	Meetings (p_{txyc})	Unobserved heterogeneity ($log(\sigma)$)	Surplus 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)

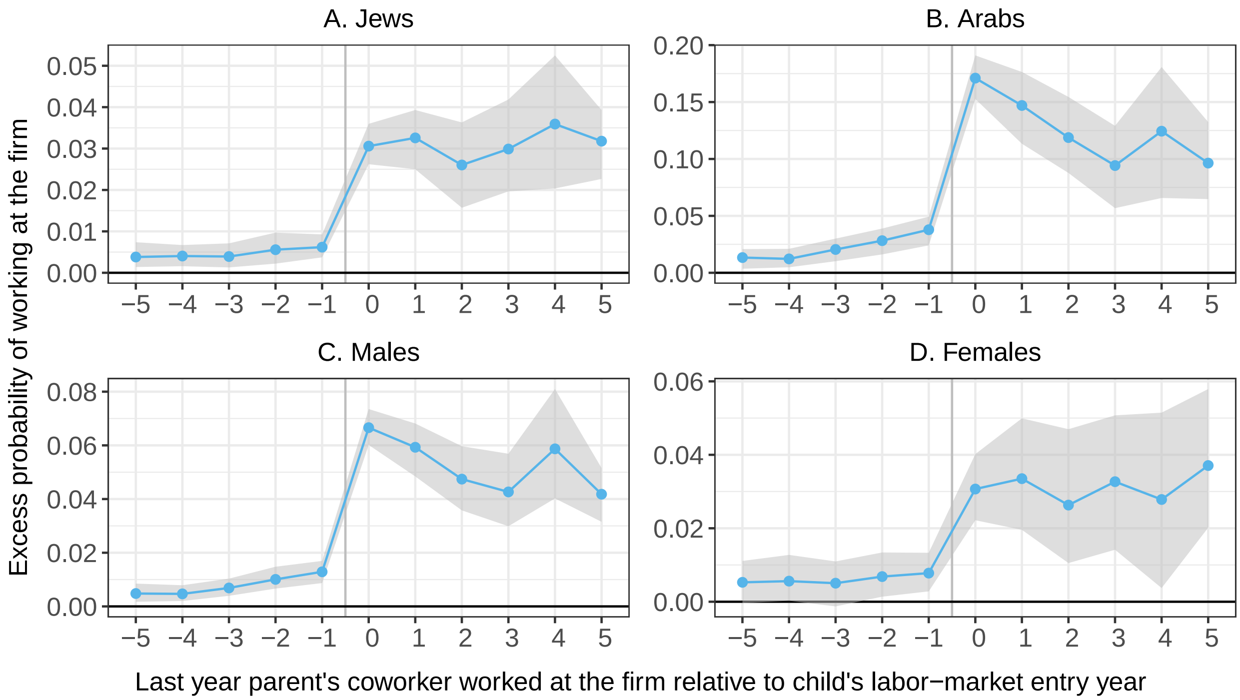
Notes: This table reports measures of the model's fit to the data (Panel A), the model's precision, and the results of Monte Carlo simulation (Panel B). The first row reports the average difference between the predicted and true moments on a logarithmic scale (averaged over all $txyc$ cells with weights 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 surplus 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 σ, 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 surplus 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.

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



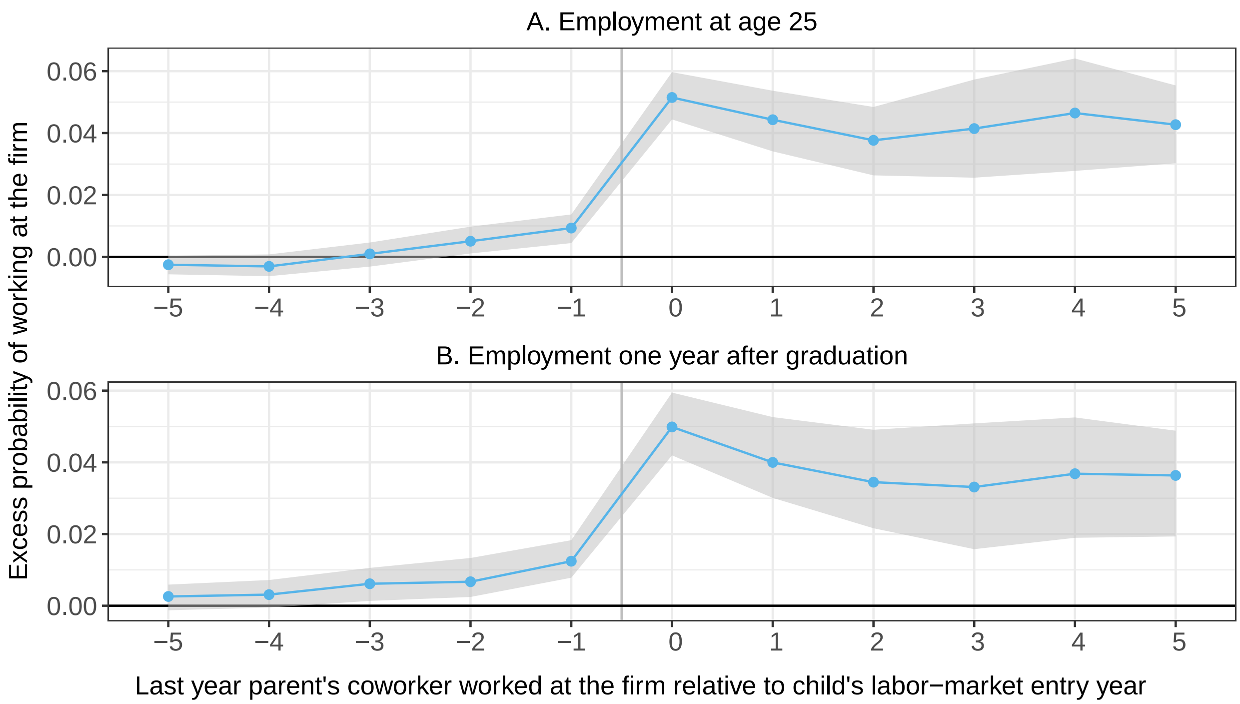
Notes: This figure shows the raw probability of working in a firm as a function of the difference between the last year the parent's coworker worked at the firm and the worker's labor-market entry year. The employment outcome is scaled by 100. The vertical line between -1 and 0 indicates the change from worker-firm pairs with phantom connections to pairs with weak connections.

Figure A2: Event-study plot of coefficients: Effect of weak parental connections on firm assignment (by groups of workers)



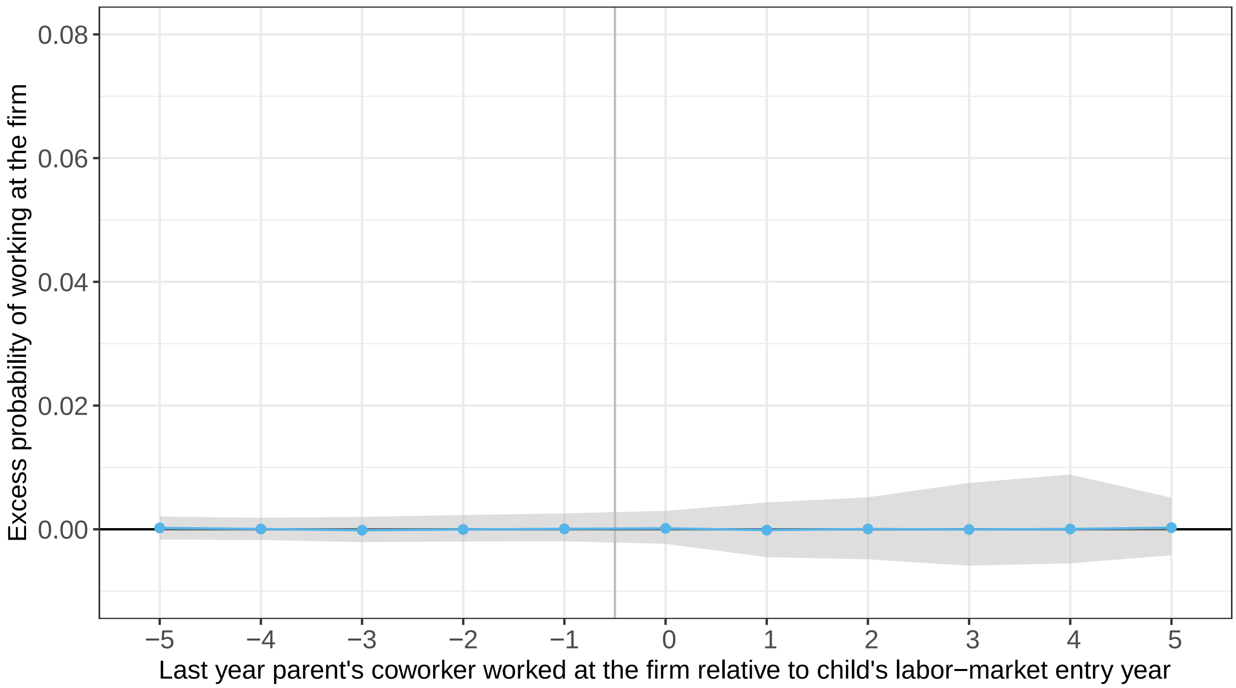
Notes: This figure shows the probability of working in a firm as a function of the difference between the last year the parent's coworker worked at the firm and the worker's labor-market entry year, relative to working in a non-connected firm, separately for sub-groups of workers. The figure shows the mean (and 95 percent confidence interval) of the estimated coefficients of phantom and weak connections across 100 estimations of equation (2) using a 20 percent random sample of workers each time. The employment outcome is scaled by 100. The vertical line between -1 and 0 indicates the change from worker-firm pairs with phantom connections to pairs with weak connections.

Figure A3: Event-study plot of coefficients: Effect of weak parental connections on firm assignment (alternative definitions of the labor-market entry year)



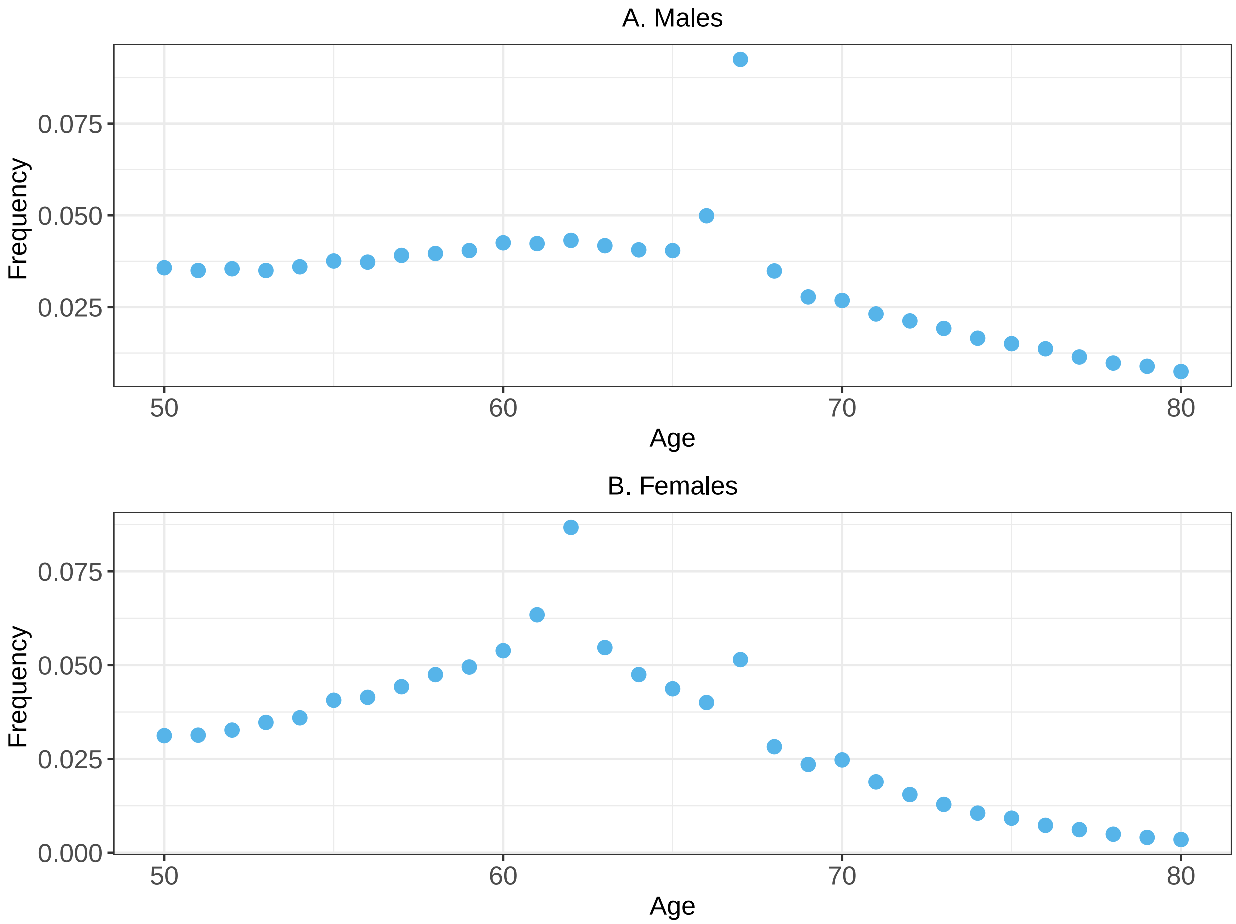
Notes: This figure shows the probability of working in a firm as a function of the difference between the last year the parent's coworker worked at the firm and the worker's labor-market entry year relative to working in a non-connected firm, for alternative definitions of the labor-market entry year. Panel A uses the year when the worker is aged 25 as the entry year, while panel B uses the year after the graduation year as the entry year. The figure shows the mean (and 95 percent confidence interval) of the estimated coefficients of phantom and weak connections across 100 estimations of equation (2) using a 20 percent random sample of workers each time. The employment outcome is scaled by 100. The vertical line between -1 and 0 indicates the change from worker-firm pairs with phantom connections to pairs with weak connections.

Figure A4: Event-study plot of coefficients: Effect of weak parental connections on firm assignment, placebo test



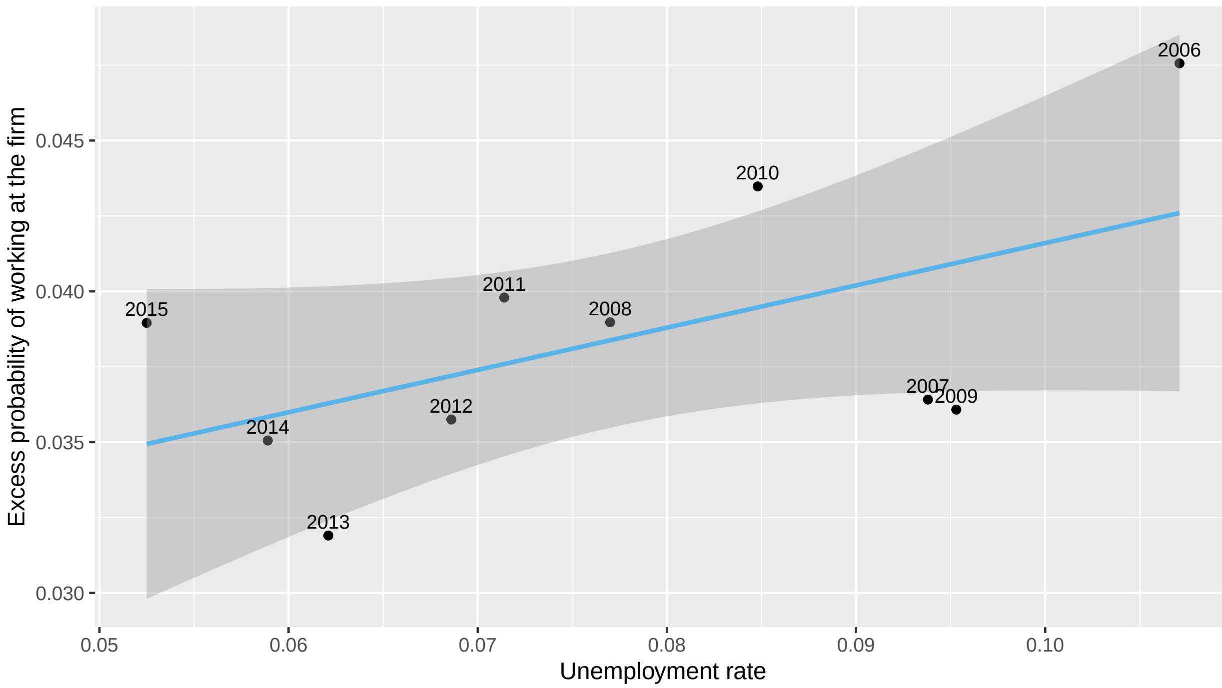
Notes: This figure reports the results of a placebo test, assigning the worker's connections to a random worker in her group. The figure shows the probability of working in a firm as a function of the difference between the last year the parent's coworker worked at the firm and the worker's labor-market 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 (2) using a 20 percent random sample of workers each time. The employment outcome is scaled by 100. 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 A5: 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 A6: Correlation between the effects of weak parental connections on firm assignment and total unemployment rate by year

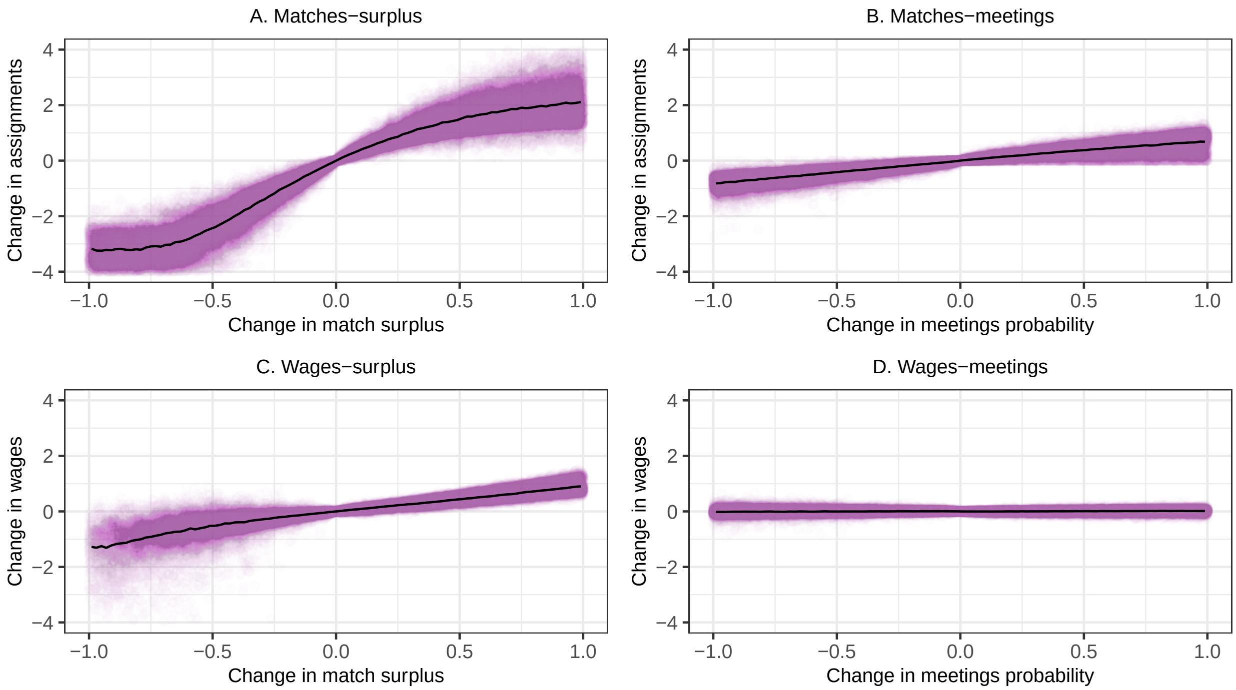


Notes: The vertical axis is the probability of working in a firm with weak connections relative to the probability of working in a phantom-connected firm for workers by labor-market entry year. The vertical axis is the total unemployment rate of the total labor force in that year. The coefficients (and robust standard errors) of the fitted line are 0.028 (0.007) and 0.14 (0.089) for the intercept and slope, respectively. The Pearson correlation coefficient is 0.55

Figure A7: Scatter plot: Lower and upper wage bounds

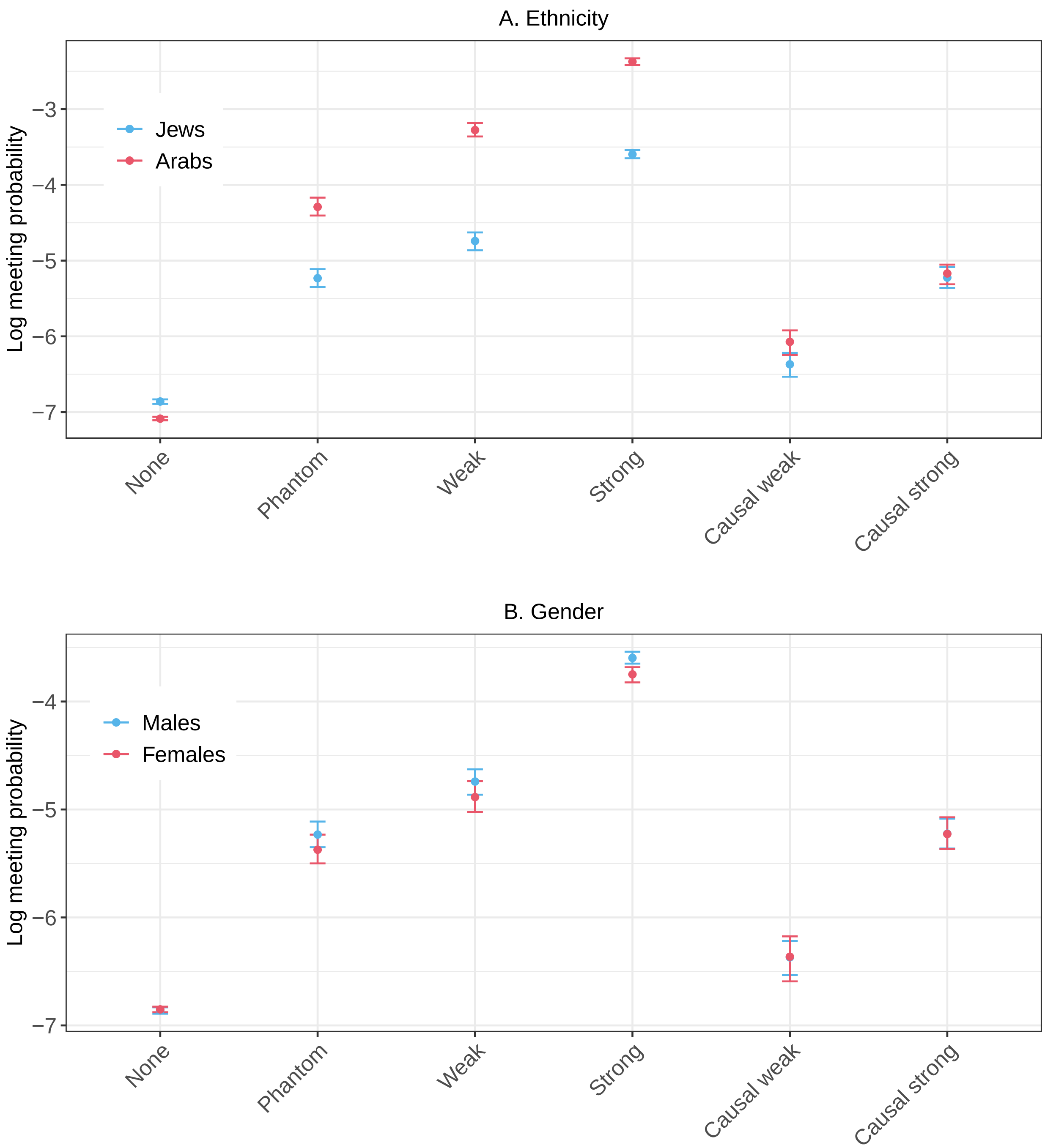
Notes: This figure shows the relationships between lower and upper wage bounds that support the equilibrium matching. The black line shows the mean value of the wage upper bounds for 100 bins of the lower bounds.

Figure A8: Scatter plot: Changes in moments as a result of changes in parameters of the same group of workers and firms



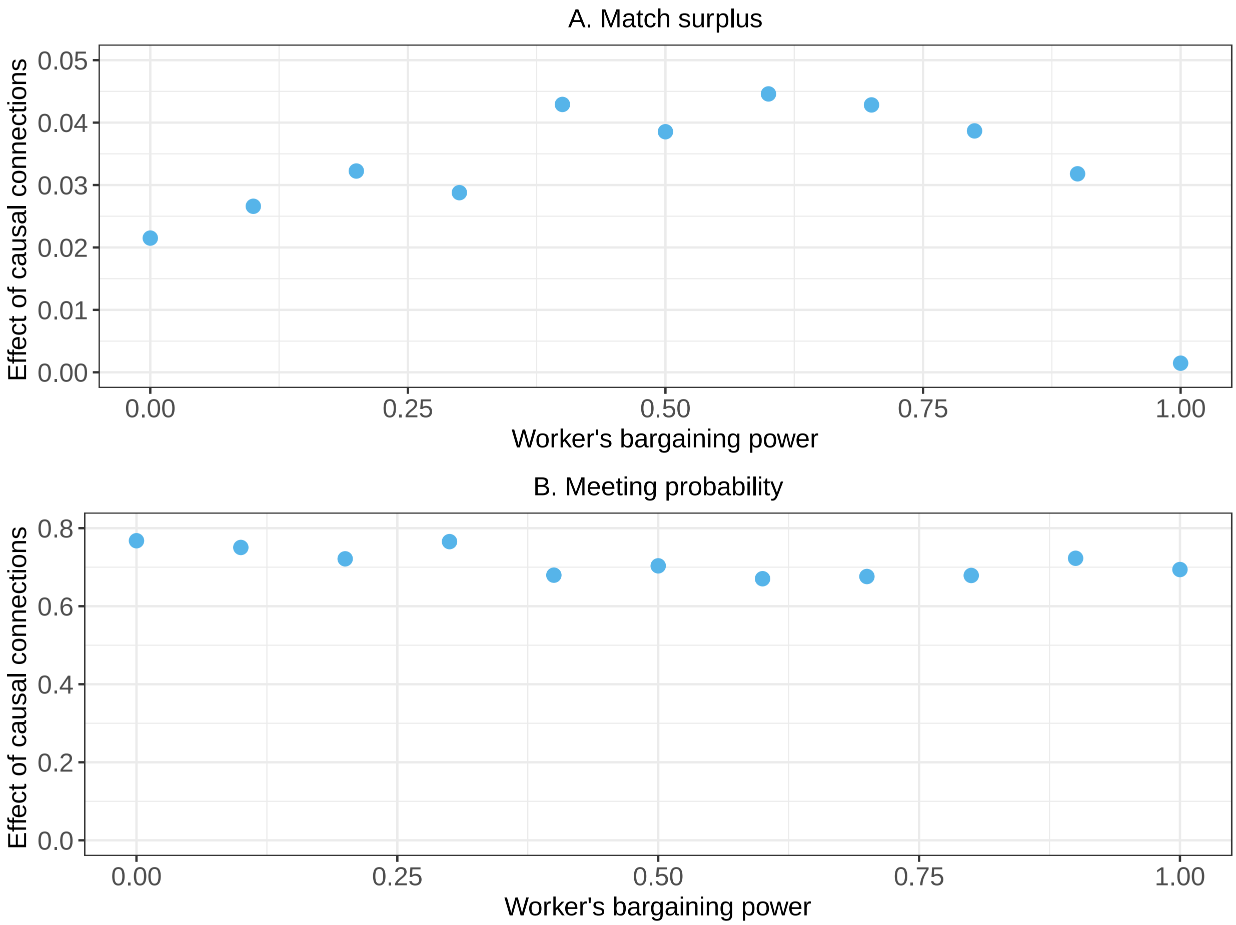
Notes: This figure shows the relationships between the parameters of the model and the predicted moments. I run 10,000 simulations of the model. Each time, I change the value of only one parameter, either the match surplus β_{txyc} or the meeting probability p_{txyc}, of one xyc group in each market 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 A9: Model estimates: Average meeting probability by workers' group and connection type



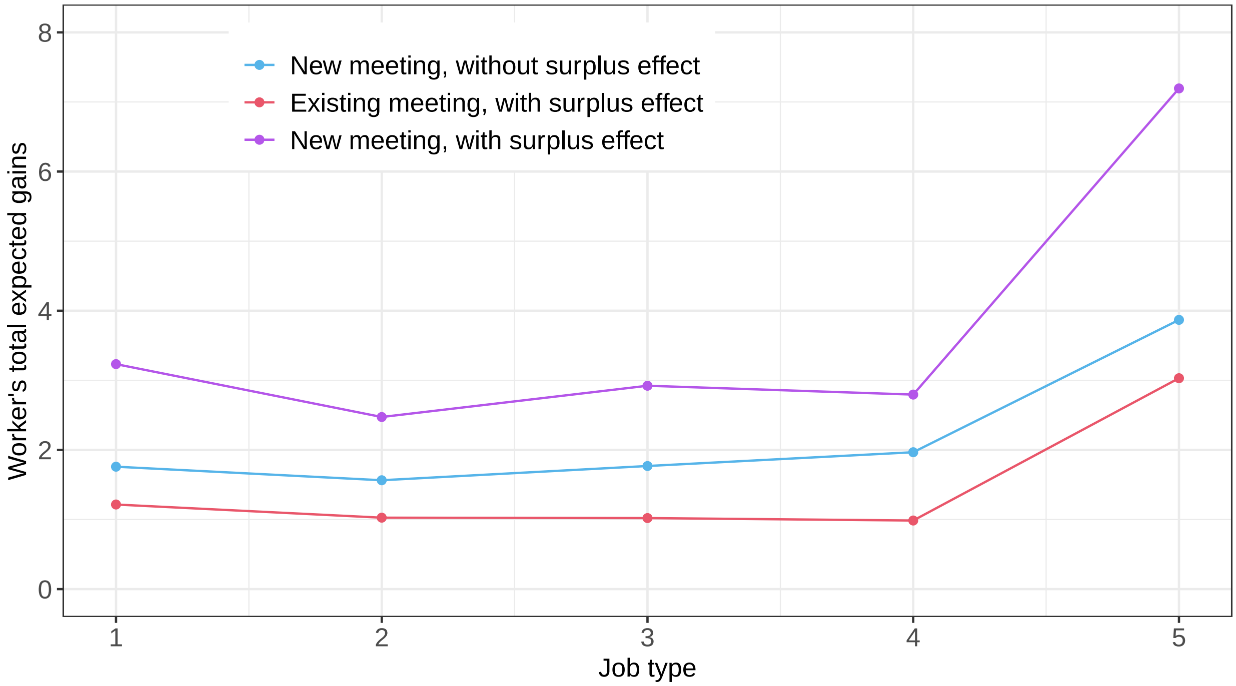
Notes: This figure shows the results of regressing the log of the meeting probabilities obtained from the model on worker, firm, and connection characteristics, and the interactions between worker and connection features. I estimate the regression using weighted least squares, with weights equal to the 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 95% confidence intervals).

Figure A10: 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 surplus 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 surplus 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.

Figure A11: 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 one of three 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 surplus to reflect the surplus of a causal weak connection, and 3) adding a random meeting with causal weak connections. The surplus of a causal weak connection is the excess surplus of weak connections compared to phantom connections. Each statistic is calculated separately for each of the 100 estimations of the model, based on 1,000 new meetings/connections for each estimation, and the table reports the averages across the 100 estimations.

B Data appendix

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

Employment and wages: The data contain observations at the worker $\times$ firm $\times$ year level. In each observation, there are monthly employment indicators and total yearly salary. I cleaned the data by: 1) dropping observations with missing worker or firm identifiers, 2) replacing empty monthly indicators with zeros, 3) dropping observations that are duplicate in all variables, 4) for duplicate worker-firm observations, taking the maximum of the monthly indicators and the sum of the yearly earnings

Parents and Children: The data include the identifiers of the mother and the father of each Israeli citizen, provided that they are also Israeli citizens.

Education: Starting in 1996, I observe the higher-education institution and period of enrollment of each individual in Israel. "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.[51]

Workers' location: I measure the new worker's residence location by the longitude and latitude of the centroid of the city she lived in at 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 with no 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

[51] According to the Israeli Central Bureau of Statistics definitions, "Others" refer to Non-Arab Christians, members of other religions, and not classified (CBS 2019). The 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 (2011 Israeli classification).

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.[52]

Panel data construction: 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 or her 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.[53] The period of the sample is 1991–2015 and I keep workers aged 22-69 each year. I construct a second dataset from this panel dataset, keeping only firms with 5-500 workers each year. I use these data to build the parental network over time and the sample of "new workers" (see below).

First stable job and labor-market entry year: 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 during a calendar year and produces total annual earnings corresponding to at least 150% the national average monthly wage.[54] Labor-market entry year is the year the new worker finds her first job.[55]

[52] The data do not include establishment/plant identifiers or an indicator for multi-plant 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.

[53] The minimum monthly salary in 2015 was 48.8% of the average salary in that year. This ratio fluctuated between 40%-50% in 1991-2015. Therefore, I exclude workers who earn approximately 50% or less the minimum wage, similarly to Kramarz and Skans (2014).

[54] I focus on the employment and salary of young people when they enter the labor market for two reasons. First, young workers' first job experiences are important for their future careers (Oreopoulos et al. 2012; Arellano-Bover 2020). Second, focusing on first-job outcomes enables isolating the impact of the initial set of connections the workers enter the labor market with—parental professional network in this case—from the connections the worker herself forms at the labor market (and might be impacted by the initial connections as well).

[55] 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. I choose not to do this in the main analysis for two reasons. First, it is challenging to differentiate people who unsuccessfully looked for a job from those who did not look for a job based on employment information alone. For example, many Israeli youths postpone their entry into the labor market because they take a long backpacking trip following

New Workers: My analysis sample of "new workers" comprises Israelis who found their first stable job between ages 22-27 in the years 2006-2015 in a 5-500 workers firm.[56] I exclude workers without any parent that worked in a 5-500 workers firm when they were 12-21 years old. I further exclude immigrants and Ultraorthodox Jews from the sample.[57]

Firms' pay premium: The firm pay premiums are estimated using the AKM model (Abowd et al. 1999):

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

where w_{it} is the log of monthly salary of worker i at year t, α_i is person fixed effect, $\psi_{J(it)}$ is firm fixed effect, $Z_{it}' =$ are set of year fixed effects and cubic age function restricted to be flat at age 40 (Card et al. 2018). 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").[58]

Weak connections: a worker i has weak connections to a firm j if i's parent and a worker k worked simultaneously at a firm $j' \neq j$ 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 employ between 5 and 500 employees.

military service (Noy and Cohen 2005). Second, studying the impact of connections on the probability of finding a job at a fixed age, or at a specified time after graduation, might bias the estimates. For example, if the worker 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 (see definition below) even though the worker had active connections there when she joined the firm. However, I checked the robustness of the results to such definitions of the labor-market entry year (see Appendix D.2).

[56] Intuitively, the probability that a random pair of workers form social connections decreases in the firm's size. In the paper, 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, firms with 5-500 workers accounted for 29.6% of the firms and employed 52.2% of the workers (Table A1).

[57] Oftentimes, immigrants do not have parental connections in the labor market. See Arellano-Bover and San (2020) for the role firms play in explaining the pay gaps between former Soviet Union immigrants and natives in Israel. Ultraorthodox Jews in Israel have unique labor-market characteristics, such as low (secular) education and employment rates, especially for males (Berman 2000; Fuchs and Epstein 2019). Specific research is needed to study this group.

[58] These premiums aim to capture the average differences in salary firms pay to similar workers. They 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.

Phantom connections: a worker i has phantom connections to a firm j if i's parent and a worker k worked simultaneously at a firm $j' \neq j$ when i was 12-21 years old, and k worked at a firm j at any time within five years before or after i's labor-market entry year, but not that year.

Strong connections: a worker i has strong connections to a firm j if at least one of the following conditions are satisfied: 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 at any time within five years before or after i's labor market entry year.[59]

C The role of firms and social networks in earnings inequality

In this appendix, I decompose the ethnic and gender pay gaps into between- and within-firm variation. I also check the correlation of these gaps with measures of connection quality.

To get the raw ethnic and gender gaps, I estimate the equation:

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

using all workers ages 22-69 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 A2 report the OLS estimates of equation (C1) without and with the firm fixed effects, respectively.

Starting with the ethnic pay gap, the overall gap between Jews and Arabs in 2015 is 25.3 log points. Controlling for firms decreases the ethnic pay gap to 5.1 log points. Comparing

[59] Two components of these definitions are noteworthy. First, 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 formed at the years between the child is 22 until the year she enters the labor market. Doing so would mechanically increase the set of connections available for workers that enter the labor market later. Second, 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 unclear 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 and parents' past firms eliminates weak connections estimates' upward bias. Three, keeping both weak and phantom connections with only one contact makes them comparable. It 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 alternative definitions of connections (see Appendix D.2).

the ethnic pay gap estimates 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 points. Controlling for firms decreases the gap to 28.8 log points. Those results indicate that, unlike the ethnic gap, most of the gender gap (78%) is explained by within-firm variation.

Table A3, 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. 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 A3, 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 A3 I add firm fixed effects to the regression. The coefficients of the correlation between parental connections and salary become 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 also report the coefficients for the gender pay gap. Table A3 shows that the gender pay gap patterns are different. First, most of the gender pay gap is explained by within-firm variation (columns 1-2). Second, including connections in the regression does not affect the magnitude of the gender pay gap (columns 1 and 3).

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.

D Regression appendix

D.1 The econometric model

Starting with equation (1), let $D_{ij} \equiv \max_c \left[D^c_{ij}, c \in \{p, w, s\}\right]$ be a variable that indicates whether a worker i has any type of connections in firm j. First, I restrict the sample under study to cases in which there is within group-firm variation in D_{ij}. This restriction has no impact on the parameters of interest since the discarded observations are uninformative conditional on the fixed effects (Kramarz and Skans 2014). I then aggregate the model by computing, for each group-firm combination, the fraction of workers with connections in the firm that this firm hired:

$$R^{CON}_{xj} = \frac{\sum_{i \in x} e_{ixj} D_{ij}}{\sum_{i \in x} D_{ij}} = \phi_{xj} + \sum_{c=p,w,s} \delta^c \cdot D^c_{xj} + \epsilon^{CON}_{xj} \tag{D1}$$

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

$$R^{-CON}_{xj} = \frac{\sum_{i \in x} e_{ixj}(1 - D_{ij})}{\sum_{i \in x}(1 - D_{ij})} = \phi_{xj} + \epsilon^{-CON}_{xj} \tag{D2}$$

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

$$R_{xj} \equiv R^{CON}_{xj} - R^{-CON}_{xj} = \sum_{c=p,w,s} \delta^c \cdot D^c_{xj} + \epsilon^R_{xj}. \tag{D3}$$

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^c_{xj} , $c \in \{p, w, s\}$ 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 (D3) measure the effect of the different types of parental connections.[60]

Even after the fixed-effects transformation, limited computational resources prevent 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

[60]Note that, by definition, $D^p_{xj} + D^w_{xj} + D^s_{xj} = 1$, which means that the independent variables in equation (D3) are collinear. However, the estimation of that equation is feasible because the regression is estimated without an intercept.

estimates obtained, I calculate the mean and 95 percent confidence intervals of the regression coefficients and the other statistics of interest.

D.2 Robustness check: changing the definitions of parental connections and labor-market entry year

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".[61] The first column of Table A4 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 weak and phantom connections coefficients 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 statistically 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 A4, 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 that 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 indicate that the estimated effects using the baseline definition of parental connections are 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 a single indirect connection 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.

Following Kramarz and Skans (2014), I define the labor-market entry year as the year the new worker finds her first job. This definition may be endogenous if social connections affect not only the identity of the firm a worker finds, but also the time until she finds a

[61] See the discussion in Appendix B.

stable job (or her decision to start looking for jobs). I check the robustness of the results to the definition of the labor-market entry year, by estimating the effect of connections using two alternative definitions: 1) The year of which the worker is 25 years old, and 2) The year after the worker's graduation year. Figure A3 plots the event-study coefficients using these definitions. Overall, the results look very similar to the benchmark results, with discrete increase in the probability of employment at time zero of about 0.04 percentage points.

D.3 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 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.[62] Column 1 of Table A6 shows the estimated coefficients. As expected, compared to firms with no connections, firms with all three types of social connections are 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 than unconnected workers. This correlation, however, is similar to phantom and weak connections, with estimates of 0.077 and 0.076 percentage points, respectively (Table A6, column 2).

D.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 is no systematic difference in the

[62] 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.

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 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. During the analysis years, the statutory retirement age in Israel is 62 for females and 67 for males. At that age, workers are entitled to leave their job and receive a pension. Figure A5 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 A7 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, respectively). 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 A7 column 1). However, due to the small number of such cases, the estimated ratio for "dead" connections is not statistically 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 are 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 left 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 A7 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 a firm with phantom connections is slightly smaller for "death" and somewhat larger for "retirement" than other connections. However, due to the 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.

D.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, suppose connections leave (become "phantoms") when demand for a particular type of labor is falling. In that case, the firms that usually hire this type of labor will hire 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 actual- and phantom-connected firms is mediated by other factors correlated with the different types of connections, the probability of a worker working in a firm that another group member has real connections to will be higher than in a firm that another group member has phantom connections to. On the other hand, if the employment probability is higher only if the connections are the worker's true connections (and not the connections of someone else with similar observable characteristics), that suggests that the estimated effect is the effect of the connections themselves.

Table A8 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 and a phantom connected firm. The event study estimates of equation (2) also showed no difference between phantom and real connected firms (Figure A4).[63]

E Model appendix

E.1 Links with theoretical models of social connections

In this section, I study whether my model is compatible with various theoretical models of social connections in the labor market. I separate all models into four sets, according to the exact mechanism in which social connections impact the matching probabilities and wages. The first two sets are nested in my model in the search frictions and match surplus mechanisms, respectively. The third sets are models that, in principle, can be included in my framework but are not included in the current study due to a lack of appropriate data. The last set of models is outside my framework.

Models where social connections reduce search frictions: I first consider models where social connections increase the probability that the job seeker and the employer are aware of the existence of each other and consider forming a match. This structure is compatible, for instance, with Calvo-Armengol and Jackson (2004) and Fontaine (2008) who assume the role of social networks is to help workers obtain information about job opportunities. Likewise, social connections can improve the firms' information flow about potential candidates. The first stage in my model captures this mechanism.

Models where social connections affect the firms' match surplus: Social connections might affect the firm's surplus from the prospective match for several reasons. First, social connections between workers in a firm might directly affect the firm output, either positively or negatively. Rotemberg (1994) shows that close relationships between coworkers may lead employees to work harder if their payment depends on their joint output. Likewise, social connections between managers and workers can help firm performance if they allow managers to provide nonmonetary incentives to workers or help reduce informational asymmetries within the firm. They can also harm the firm performance if managers display favoritism toward workers they are socially connected with (Bandiera et al. 2009).

Second, social connections may reduce uncertainty about the productivity of the worker or the match. This information transmission mechanism might result from employees explicitly

[63] The fact that the estimated effect is not different from zero for the phantom-placebo connections is expected because the control group ("no connections") only includes worker-firm pairs in which someone from the worker's observable group has some type of connections in the firm (see the discussion after equation (1)). Still, if active (weak) links reflected a more positive employment trend in a firm than phantom links, we would see an increase in the employment probability for active connections (of someone else).

recommending network members to their employers and thereby provide them with information about potential job market candidates that they otherwise would not have (Dustmann et al. 2016; Bolte et al. 2020). Likewise, even without explicit recommendation, current employees are likely to have network members of similar quality given assortative matching in personal networks. They hence provide information about the quality of the candidate (Rees 1966; Montgomery 1991). Either way, because hiring, training, and firing workers are costly, lower uncertainty about the match quality positively impacts the firm value of the match.

Finally, social connections might impact the hiring decision and wages due to favoritism (Beaman and Magruder 2012; Dickinson et al. 2018). My model captures this mechanism by assuming that the firm, or workers that make the hiring decision within the firm, get higher utility from hiring connected workers.

Models that can be included in my framework given appropriate data: Social connections in the workplace might also have an impact on the utility the worker receives from the match. The presence of a friend in the plant may be an important "fringe benefit", making the job more attractive to the worker. Alternatively, social connections might provide the worker with better information about the job (e.g., fairness of supervision in a factory) (Rees 1966).[64] The model estimated in this paper ignores non-wage benefits that might impact the workers' utility, implicitly assuming no systematic differences in workers' utility between connected and non-connected jobs (besides the impact of connections on wages).[65] However, with additional data such as workers' satisfaction, productivity, applications, etc., it is possible to use the model and estimating method proposed here to estimate this mechanism.

Likewise, social connections in a firm might affect the workers' bargaining power. It might be because they have better information on the salary they should ask for or other types of inside information.[66] Also, due to nepotism, the firm might want to pay a higher wage to a connected worker, which, again, can be captured by higher bargaining power (i.e., a higher share of the surplus the worker gets). The current model assumes a common bargaining power that is not a function of social connections but can be extended to include the impact of connections on bargaining power given additional data.[67]

[64]Social connections might also increases the rate of on-the-job social network and human capital formation (see Bonhomme et al. (2019) and Arellano-Bover and Saltiel (2021)). These effects can be implicitly captured in this static framework by a higher (non-wage) utility the worker gets from working at this job.

[65]Galichon and Salanié (2020) offer to use data on transfers (in addition to the matching data) to separately estimate the agents' utility on the two sides of the market (e.g., firms and workers). However, they do not include search frictions in their model.

[66]See Roussille (2020) for the importance of information about the potential salary during the bargaining process.

[67]The two mechanisms that I do include in the estimation of the model (search frictions and firms' surplus)

Outside my framework: My model assumes that the only role of wage is to clear the market and that there is no other way to pay workers besides the current salary. However, firms may pay higher wages to connected workers to make them work harder. Likewise, firms may attract connected workers by providing them other types of benefits besides current salaries, such as a promise for a higher rate of wage growth. My model cannot capture these effects of social connections.

E.2 Finding the equilibrium matching

Given the joint surplus f_{ij} and meetings m_{ij}, the equilibrium matching can be found using the auction algorithm (Bertsekas 1998).

Definition 2 (the auction algorithm).

1. Start with an empty assignment S, a vector of initial wages w_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.,

   $$i_j = \arg\max_{i \in m(j)} f_{ij} - w_i \tag{E1}$$

   and the corresponding value

   $$v_j = \max_{i \in m(j)} f_{ij} - w_i \tag{E2}$$

   and find the best value offered by workers other than i_j

   $$q_j = \max_{i \in m(j), i \neq i_j} f_{ij} - w_i \tag{E3}$$

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

   $$b_{ij} = w_{i_j} + v_j - q_j + \epsilon \tag{E4}$$

are much more prevalent in the theoretical and empirical literature than the mechanisms I do not include (workers' surplus and bargaining power). Although it seems plausible to assume that the first two mechanisms are indeed of first order compared to the others, checking it empirically is an important direction for future research.

(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 w_i to the highest bid:

$$w_i = \max_{j \in B(i)} b_{ij} \tag{E5}$$

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 $I_t \cdot \hat{\epsilon}$ is deemed sufficiently small. Each application of the algorithm provides good initial wages 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).

E.3 Finding the equilibrium wages

Given the equilibrium matching, the bounds on the equilibrium wages can be found using the Bellman-Ford algorithm (Ahyja et al. 1993; Bonnet et al. 2018).

Definition 3 (the Bellman-Ford algorithm).

Let w_i and v_j be the equilibrium payoffs for workers and firms, respectively, in a connected set G, where the first worker's wage is normalized to zero $w_1 = 0$. The firm-optimal equilibrium wages are the fixed point of the mapping:

$$w_i = \max(w_i, \max_{j \in m(i)} (f_{ij} - v_j)) \quad , \quad v_j = \min(v_j, f_{i^*(j)j} - w_{i^*(j)}) \quad , \quad w_1 = 0 \tag{E6}$$

where $i^*(j)$ denotes the equilibrium match of firm j. The fixed point of this map can be computed by iterating on (E6) from the initial values $\{w_i = -\infty, w_1 = 0; v_j = \infty\}$. Similarly, the worker-optimal equilibrium wages can be computed by iterating on:

$$v_j = \max(v_j, \max_{i \in m(j)} (f_{ij} - w_i)) \quad , \quad w_i = \min(w_i, f_{ij(i)} - v_{j(i)}) \quad , \quad w_1 = 0 \tag{E7}$$

from the initial values $\{w_i = \infty, w_1 = 0; v_j = -\infty\}$.

Definition 4 (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 let us normalize the first worker's wages in each subgraph to zero. Then, the 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 $\{\underline{w}_i\}_{i=1}^{I_t}$ be the firm optimal wages. Assume by contradiction that $\underline{w}_i = -\infty$ 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 $\underline{w}_i = -\infty$, there exist small enough w_j such that $w_j < f_{ij} - v_j$. But this contradicts the optimality of the match. The symmetric argument holds for the worker optimal wages.

Now, assume G_t is not double connected. WLOG, assume there exists a worker i such that $|m(i)| = 1$. Assume by contradiction that $\underline{w}_i$ is finite. Let (μ, w) be an equilibrium outcome. Changing only the wage of i to $w_i = \underline{w}_i - 1$ 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 $Per^2(i)$.[68]

E.4 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 $\theta^2 \neq \theta^1$. First, 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 (Salanié 2015; Galichon and Salanié 2020).

Second, assume that p and β are unknown and only σ is known. In this case, using the information on matches only without wage data, one cannot separately identify the

[68] 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 reduced 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 value.

two underlying parameters of the model, namely the meeting probability and match surplus parameters. 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 surplus of those matches is high. However, the two parameters can be separately identified using both matches and wage data. The reason is that the two sets of moments, namely the groups' number of matches and wages, react differently to changes in the meeting rate and surplus parameters. The group's match surplus significantly impacts both the groups' number of matches and wages. In contrast, the group's meeting rate has a significant impact on the number of matches but (almost) no impact on wages.

To see the intuition for this, 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 intervention, 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 (100 for each of the 100 sets of estimated model parameters). 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 A8 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 group's meeting probability and match surplus 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 surplus parameter (Panel C). However, a change in the meeting parameter has little impact on wages (Panel D).

Table A9 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-surplus, matches-meetings, and wages-surplus 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 A9. 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 surplus 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_{txyc}, we have $\hat{\mu}^1_{txyc} \neq \hat{\mu}^2_{txyc}$. Next, assume WLOG that $\beta^2_{txyc} > \beta^1_{txyc}$. Because the number of matches is increasing in β, it must be the case that $p^2_{txyc} < p^1_{txyc}$. But because the wages are (almost) not impacted by p, this implies $w^2_{txyc} > w^1_{txyc}$.[69]

Third, identification of σ comes, again, from the fact that we observe wages. If wages are not observed, only the ratio between the match systematic surplus and the idiosyncratic surplus is identified using matches 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 search friction. But these two channels cannot be separately identified based on the observed amount of sorting. To see it, 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 meeting probability of the first $txyc$ cell in each market to a fixed level corresponding to 20 meetings per worker.[70]

In Appendix E.6, I support the informal identification arguments with Monte Carlo simulation.

[69] 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.

[70] 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 workers and firms several times. I do not explore this in the current research.

E.5 Estimation and counterfactuals

Moments: I estimate the model using two sets of moments at the $txyc$-cell level: 1) the number of matches μ_{txyc}, and 2) the average wage w_{txyc}. I also use the within-group and overall wage variance. I calculate the residuals of the wages controlling for groups of year by age, and then add the overall mean wage. In addition to these moments, I also use the number of connections d_{txyc} in each cell in the estimation (see below).

I performed the estimation of the model outside the National Insurance Institute's research laboratory. To ensure data security, the National Insurance Institute prevents the export of any information for groups of less than ten individuals. Therefore, I do not use matches and wage information on $txyc$ cells with less than ten 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 ten matches, corresponding to less than 1.5% of the workers (and jobs).[71]

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 surplus shocks ξ_{ij} from a standard normal distribution.

Next, 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 $\bar{p}_0 = M * T/I$, with $M = 20$ meetings per worker on average.[72]

Empty cells: To allow the possibility of $txyc$ cells with no matches, in the estimation equations (13) and (14), I calculate $log(z+1)$ instead of $log(z)$. In equation (14), the average

[71] Because all the results I report are weighted by the number of workers/jobs in each cell, the potential bias of excluding those cells is limited.

[72] Using this normalization, I get 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).

wage of a cell w_n is multiplied by the number of matches in the cell. Therefore, there is no need to know the average wage of cells, only the total wage, which allows the inclusion of empty cells in the analysis.

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 this cell's meeting or surplus parameter will also lead to the same predicted moments. I address this problem in two ways. First, when calculating aggregate statistics and results, such as the average impact of connections on the meeting and surplus parameters, I weight each observation by the observed number of matches, which gives no weight to empty cells. Second, when calculating the "causal" connection parameters in the counterfactual exercise, I censor the top and bottom 1% of outliers, weighted by the number of observations.

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 iteration in any of the 100 simulations. 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:

$$log(\mu_{txyc}/d_{txyc}) = a + p_c + \epsilon_{txyc} \tag{E8}$$

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 $p^0_{txyc} = \bar{p}_0 \cdot \hat{p}_c$, where $\bar{p}_0$ is the normalization level of the meeting parameter described above.

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

$$log(w_{txyc}) = b + \phi_1 Arab_x + \phi_2 Educ_x + \phi_3 Female_x + \psi_y + \delta_c + \epsilon_{txyc}, \tag{E9}$$

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.

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 square difference between actual and predicted (log) moments (averaged across $txyc$ cells if applicable) of one of the four sets of moments ($\mu_{txyc}, w_{txyc}, Var_w, WithinVar_w$) in N_{tol} iterations in a row. I use $\epsilon_{tol} = 10^{-10}$ and $N_{tol} = 50$.

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

Causal connections: The surplus parameters of causal connections are calculated as the excess impact of real connections and compared to phantom connections (see equations 18-19). As mentioned above, the estimated accuracy is low for cells with a small number of observations. To account for that and to avoid extreme values, I censor the top and bottom 1 percent of the parameter estimates, weighted by the number of observations.

E.6 Model fit and precision, and Monte Carlo simulation

Panel A of Table A10 reports measures of the fit of the model 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, respectively. The predicted wage variance and within-group wage variance are also 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 0.998 for the average wage. Overall, Panel A of Table A10 shows that the model fits the data well, which means that the update mapping successfully inverts the information on the moments into the parameters.[73]

The precision of the estimates is also high. Panel B of table A10 compares the the model's 100 sets of estimated parameters. The first row reports the average correlation in the surplus and meeting parameters across any possible pair within the 100 sets of estimated parameters. The average correlation is 0.980 for the surplus parameter and 0.988 for the meetings parameter. To check the precision of the unobserved heterogeneity, σ, and the surplus-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, respectively).

Finally, I investigate the identification of the model by Monte Carlo simulation. I generate data using the model, assuming the average parameter values described above are the "true" parameters. 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

[73]This result does not say 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 shows that the algorithm successfully inverts the data, although I do not have formal theoretical results to guarantee it.

average correlation between each set of Monte Carlo estimates and the "true" parameters is 0.972 and 0.985 for the surplus and meeting parameters, respectively (Table A10, Panel B, third row). The average estimated unobserved heterogeneity and surplus scale are -1.076 and 9.186, which are also close to the "true" parameters, -1.069 and 9.174, respectively (Table A10, Panel B, fourth row). Overall, the results of the Monte Carlo simulation suggest that the proposed estimation procedure can identify the true parameters of the model.

E.7 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 A10 plots the difference between the average estimated effects of weak connections and phantom connections on the surplus and meeting parameters for different workers' bargaining power values. Starting with the match surplus parameter, the estimated effects of causal weak connections (the difference between the effects of weak and phantom connections) are always positive. They 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.[74] The only exception is the unrealistic scenario that workers have perfect bargaining power. In this case, the estimated effect is close to zero (Figure A10, Panel A).

Likewise, the estimated causal effects of weak connections on the surplus 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 A10, Panel B).

[74]The value in the benchmark model is the average across 100 different sets of estimated parameters of the model with $\lambda = 0.5$, whereas in Figure A10 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.