Social Connections and the Sorting of Workers to Firms*

Marcus Eliason[a,d], Lena Hensvik[a,d], Francis Kramarz[b,c],
Oskar Nordström Skans[c,d,a]

[a]*Institute for Labour Market and Education Policy Evaluation (IFAU), Sweden;*

[b]*CREST, France;*

[c]*Department of Economics, Uppsala University, Sweden;*

[d]*Uppsala Center for Labor Studies (UCLS), Sweden*

Abstract

We document how searching for jobs using social networks affects the allocation of workers to new positions. We assess the presumption that social networks reinforces unequal access of high-wage workers to high-wage establishments. Our results based on very detailed Swedish register data contradict this view. Although high-wage job seekers tend to be connected to high-wage workers employed in high-wage establishments, and although social connections directly cause the allocation of workers to jobs, we find that low-wage firms are more likely to use social connections to hire high-wage workers. Overall, sorting resulting from market hires is *more* unequal than sorting resulting from social networks hires.

Keywords: networks; job search; job displacement; hiring

JEL classification: J60; J30; J23

*We are grateful for comments from Pierre Cahuc, François Fontaine, Giuseppe Moscarini, Rafael Lalive, Xavier D'Haultfoeuille, Magne Mogstad, and seminar audiences at Yale, LSE, Lausanne, Mainz, Oslo, IFN, CREST, SOLE 2014, ESSLE 2013, SOFI, EEG-NIPE, Århus, Uppsala University, SSE, 9th Nordic Workshop on Register Data and Economic Modelling, Annual Symposium in LE (Rome 2019), the Warwick Labour Workshop, and the Models of Linked Employer-Employee Data conference (NY, 2019). Kramarz and Skans acknowledge funding from the ERC advanced grant FIRMNET (project ID 741467). Hensvik and Skans acknowledge funding from the Ragnar Söderberg foundation.

1 Introduction

A very large literature has identified social networks as a key mediator in the process of matching workers to firms. Even though access to networks is often perceived as unequal, the causal effect of social networks on the allocation of workers to jobs has rarely been analyzed empirically. This lack of evidence is particularly glaring in the context of *sorting inequality* – defined as inequality arising from the sorting of workers to firms. In this paper, we present what we believe to be the first comprehensive empirical analysis of the effects of social networks onto differential access to jobs and the ensuing effects on sorting inequality.

Our question is motivated by a series of recent and influential studies showing that positive sorting in the matching between workers and firms affects the wage structure. Using a statistical decomposition of wages into fixed person and firm effects in the spirit of Abowd et al. (1999), AKM hereafter, (Card et al., 2013) has shown how the entry of low-paying firms in Germany affected the overall wage distribution and how sorting of men and women into different firms in Portugal is related to gender wage disparities (Card et al., 2016). These studies show that sorting inequality arises from high-wage (resp. low-wage) workers' disproportional access to jobs in higher (resp. lower) paying and more (resp. less) productive firms, a result corroborated using alternative methods and/or data by Barth et al. (2016), Bonhomme et al. (2018), Abowd et al. (2018), Card et al. (2018), and Song et al. (2019) among others.[1] An important contribution, related to ours, is Schmutte (2015) who also uses the AKM-framework to show that neighborhood networks provide workers with access to high-wage firms.

Across the social sciences, researchers often view social networks as an important source of unequal access to jobs. The following quote by an authority in the field illustrates this tendency very clearly:

> "[...] Social networking, which claims to make connections and bring people together, paradoxically exacerbates social divisions and inequalities. Social networks are inherently unfair and exclusionary. They operate on the principle [...] 'birds of a feather flock together.' [...] If people have lower prestige, socio-economic status, or are the targets of discrimination, then their networks will be composed of people with lower prestige, lower socio-economic status, and who are otherwise disadvantaged." (Kadushin, 2012).

The economic literature on job search networks is a case in point. Widely cited theoretical articles relate social networks and labor market inequality through the preferential dissemination of information about vacancies (Calvó-Armengol and Jackson, 2004) and referral opportunities within social networks

[1] A related set of studies relies on structural estimation of models with on-the-job search following in the tradition of Postel-Vinay and Robin (2002). In this structural literature, the focus is on the job-to-job mobility process where productive workers receive and accept offers from more productive firms through job ladders. But as in the reduced form literature, the role of social networks is rarely explored.

(Montgomery, 1991). Furthermore, various job search models (see Ioannides and Datcher Loury (2004)) appear to imply that strong social ties induce more unequal access to jobs than weaker ties.

Because social networks play a quantitatively important role in the process of matching workers to jobs – between one-third and one-half of all job matches are usually attributed to social connections (see, e.g., Ioannides and Datcher Loury, 2004) – studying the mechanisms at work when social connections are involved in the matching process is likely to shed light on key determinants of labor market sorting in general.[2] Social connections help alleviate some of the information problems that agents face when searching on a frictional labor market. Connections help inform agents (usually workers) about the presence of available opportunities on the opposite side of the market (see, e.g., Calvó-Armengol and Jackson, 2004, 2007; Calvó-Armengol et al., 2007) or help inform agents on one side of the market (i.e., the firms and/or the workers) about the properties of agents on the opposite side (see, e.g., Montgomery, 1991; Simon and Warner, 1992; Casella and Hanaki, 2006; Dustmann et al., 2015; Galenianos, 2013).[3] Social connections may also help inform workers about idiosyncratic job amenities which may be important for sorting (Card et al., 2018), or serve as an independent source of amenities if workers prefer to work with people they know as in Heath (2018).

In this paper, we assess the contribution of social networks to the unequal access to jobs because of their role in the sorting of workers to employers. We rely on three building blocks:

1. data on the structure of workers' social networks as inferred from register data on social connections;

2. estimated person and establishment fixed-effects from an AKM decomposition;

3. establishment closures as events forcing groups of displaced workers to search for new jobs under similar conditions;

For the first building block, we use Swedish longitudinal administrative data to measure multiple types of social relations for every individual; we study family members, former co-workers, former classmates, and current neighbors.[4]

For the second block, we characterize workers and employers through estimates derived from the AKM model. In contrast to much of the literature, e.g. the recent studies by Card et al. (2013) and

[2] As argued by Galenianos (2013), both information channels are closely related to the use of an aggregate matching function to approximate search frictions. Oyer and Schaefer (2016) make a similar argument from the personnel-economics perspective. They argue that we know too little about the strategies used by firms when filling jobs, explicitly mentioning the use of referrals as a key strategy.

[3] For related empirical studies investigating the information content of connections see, e.g., Brown et al. (2016), Burks et al. (2015) and Hensvik and Skans (2016).

[4] Despite the well-documented role of specific social networks in the allocation of workers to jobs (e.g., former co-workers, ethnic or university/alumni networks, and neighbors), most studies have examined each type of network in isolation: neighbors (e.g., Bayer et al., 2008; Schmutte, 2015), former co-workers (e.g., Cingano and Rosolia, 2012), immigrants (e.g., Dustmann et al., 2016), and parents (e.g., Kramarz and Skans, 2014).

Song et al. (2019), we explicitly focus on newly hired workers as we are interested in the role of social connections at the time when matches are formed. We take careful steps to ensure that our key results are not confounded by biases that may arise from the simultaneous estimation of person and firm effects in the AKM-setting (the so-called limited mobility bias, in particular), or from reverse causality if the use of connections affects wages. In the main analysis we use person effects that are estimated from pre-transition data only and relate these to establishment effects of future potential employers. This ensures that the estimation errors are independent as the estimates are drawn from different samples.[5] We further verify that our main conclusions hold if we instead classify the employers through the clustering procedure of Bonhomme et al. (2018) and/or if we instead rank firms by value-added per worker.

For the third building block, to help identify the impact of connections on hiring patterns, we focus on workers who get displaced after a firm closure. Importantly, however, we use this set-up with a very different purpose than the standard mass-layoff literature where the focus is to compare displaced workers to non-displaced workers.[6] Here, we instead let the setting provide us with a set of workers who move under similar conditions, where staying in the origin firm is no longer an option.[7] This allows us to compare multiple workers who are forced to leave the same establishment at the same point in time, but with different social networks. The strategy thus makes it possible to construct a control group who all move for exogenous reasons under similar market conditions which helps us to identify the impact of social networks on sorting.[8] To assess whether the patterns we document are valid beyond this specific setting, we provide a set of comparable estimates for the complete sample that also include endogenous movers.

By contrasting the rehiring patterns of workers from the same closure event, endowed with different sets of *pre-determined* social connections to different employers we measure how the effects of connections vary with the type of social connection, and with combinations of (displaced) person and

[5] Our estimated person effects will exclusively rely on the jobs preceding displacement. Estimation errors for these person effects are only affected by estimation errors related to their pre-displacement jobs. Estimation errors in these person effects are therefore, by construction, unaffected by errors in the estimated establishment effects of potential future employers. Our approach therefore allows us to overcome the direct impact of correlated estimation errors discussed in, e.g., Andrews et al. (2008).

[6] A large number of studies have used firm or establishment closures as quasi-experiments where workers lose their jobs for an exogenous reason. Examples of outcomes of these job displacements that have been studied are earnings (e.g., Eliason and Storrie, 2006; Hijzen et al., 2010), family income (e.g., Eliason, 2011), mortality (e.g., Eliason and Storrie, 2009*a*; Browning and Heinesen, 2012), morbidity (e.g., Eliason and Storrie, 2009*b*, 2010; Browning and Heinesen, 2012), divorce (e.g., Rege et al., 2007; Charles and Stephens, 2004; Eliason, 2012), fertility (e.g., Huttunen and Kellokumpu, 2016; Del Bono et al., 2015), alcohol-related morbidity and mortality (e.g., Eliason, 2014), children's school performance (e.g., Oreopoulos et al., 2008; Coelli, 2011; Rege et al., 2011), and criminality (e.g., Rege et al., 2009).

[7] In this respect, our approach is similar to that of recent papers by Cingano and Rosolia (2012), Glitz (2017), and Saygin et al. (2014), who all studied the relationship between the displaced workers' network quality (i.e., employment rate) and the speed of reemployment.

[8] Our focus on displaced workers is one important difference with Schmutte (2015) who studies how workers employed in high-paying firms connect their neighbors to such jobs and help them climb the job ladder. Using the timing of entry in jobs and the block structure of neighborhoods, neighbors employed in high-wage firms are rejoined by other neighbors, potentially because information gets transmitted within neighborhoods. In contrast, we study the structure of workers' various social networks, estimate their effects on hiring, and describe the resulting impact on sorting inequality.

(connected) establishment fixed effects. To reinforce the causal interpretation of our results, we perform various robustness checks and specification tests. In particular, we use cases of repeated displacements of the same worker to remove all time-invariant preferences of a given worker for specific establishments (as in the monopsony model of Card et al. (2018)). This alternative approach is identified from variation in each worker's network structure between her displacement events. Furthermore, although most workers are only displaced once, they tend to be endowed with connections to multiple establishments during this event, allowing us to use *within-worker* variation to identify the *relative* importance of connections to high- vs. low-wage employers for a given worker during a specific displacement event. As a complement to this analysis, we describe the sorting patterns of all job-to-job movers (i.e. also those unaffected by establishment closures) and compare these estimates to the corresponding estimates for the sample of displaced workers.

Our results can be summarized as follows.

First, social connections exhibit *baseline homophily*, i.e., positive sorting, in terms of earnings capacity; high-wage workers are more likely to be connected to other high-wage workers, and these high-wage workers are more likely to work for high-wage employers. It is important to note that we are measuring pre-determined sets of connections arising from where people work, who their family members are, who they went to school with and who they live close to. The networks we measure are therefore not a reflection of endogenous choices of whom each person interacts with among the different social arenas. We find that the sorting is most pronounced for professional ties, in particular past co-workers and classmates from university, whereas networks of connections within families, neighborhoods, and high schools exhibit much less baseline "homophily" in the AKM-dimensions.

Second, these pre-determined social networks predict hiring patterns. Displaced workers are much more likely to find jobs in the exact establishment where they have a social connection than other workers who were displaced in the same event, but without such a connection to that particular employer. A very conservative baseline for how large any remaining biases is provided by re-estimating our models for other very similar establishments (most notably, establishment within the same firm and municipality), finding estimates that are at most one tenth in magnitude.[9] Finally, and perhaps most importantly, we use the limited sample of workers who experience multiple displacements, to show that displaced workers are much more likely to be hired by an employer during a displacement year when she is connected to that employer relative to a displacement year when no such connection – between the same worker and establishment – is observed in our data.

Third, we show consistent evidence that a connection that is likely to be strong (e.g. connections of longer duration, more recently established connections, and connections fostered in smaller groups)

[9]The baseline is conservative since it is likely that workers' connections also may affect the hiring probability at the firm-level, not only the establishment level.

in the Granovetter (1973) sense, is associated with a larger estimated effect than a weaker tie.[10] The fact that the estimates are largest in cases where we are most confident that the agents are involved in social interactions lends further support to the network-interpretation of our estimates. Patterns across types of connections also reveal larger estimates in dimensions where it is plausible that interactions are more prevalent: the estimated causal effects of a family member is the largest, followed by that of a past co-worker, whereas a former classmate or a current neighbor entails a weaker, yet positive, effect.[11]

Fourth, the impact of social connections are much larger for low-wage *employers* than for high-wage employers, *regardless* of whether the connected workers are low- or high-wage. These patterns are unchanged if we instead sort establishments into clusters based on their earnings distribution and rank clusters by their mean wage as in Bonhomme et al. (2018), or if we rank establishments based on firm-level value added per worker. Indeed, the finding that connections matter more in less productive and low-paying employers, regardless of the workers' earnings-potential is *robust* across the methods used. Furthermore, displaced workers (of all types) with social connections to both a low-wage establishment and high-wage establishment are more likely to enter the former than the latter. This, again, holds equally for low- and high-wage workers. On average, connections have similar effects on hiring for low- and high-wage workers.[12] Throughout, our data do not support the notion that social connections should be more likely to result in a hire when they link high-wage workers to high-wage employers.

These results show that there are two opposing forces at play when social networks affect job search and, by consequence, sorting inequality. On the one hand, the pre-existing network structure exhibit baseline homophily – high wage workers are connected to high wage workers and employers – i.e. "birds of a feather flock together". And, these connections clearly matter for the allocation of workers across establishments. On the other hand, connections cause more hires when employers are of the low-wage/low-productive types, regardless of the type of worker involved – causal effects are independent of whether the "birds" (workers vs employers) have similar feathers or not. To assess what the net impact of these opposing forces are, we relate the observed (endogenous) sorting patterns of all new hires to indicators of whether the hire involved a social connection or not. The results show that the ensuing levels of sorting inequality is lower for those matches formed through social connections than for unconnected "market" matches. We show that this conclusion is valid for the set of displaced workers,

[10]Because so few studies simultaneously examine multiple types of connections, the role of ties' strength, as originally proposed by Granovetter (1973), and the associated "strength of weak ties" hypothesis (i.e., acquaintances matter more than close friends or family members) has rarely been tested, exceptions are Kramarz and Skans (2014) and Gee et al. (2017), and never in the context of sorting inequality.

[11]Furthermore, our two first key results jointly imply that the connections that exhibit least baseline homophily (family), also have the largest estimated effects. If the estimates only captured sorting on unobservables, we would instead have found the largest estimates for connections where the homophily was the strongest. This relative ordering is retained when comparing across connections *within a given worker*, ensuring that worker heterogeneity is not driving the results.

[12]This holds for all connections except for spousal ties that are used more by women (who, on average have lower person effects) than by men.

but also for the complete set of all job-to-job movers.

To summarize, our results clearly highlight the large and systematic impact that social networks have on the sorting of workers to jobs. However, our results do not support the common presumption that social homophily leads to increased inequality. Instead, our empirical results illustrate how social connections may counteract baseline homophily, resulting in more unequal allocations through market mechanisms than through social networks.

The rest of the paper is structured as follows: Section 2 describes the various components of our empirical strategy. Section 3 presents the data sources, including how we define and identify the establishment closures, displaced workers, social connections but also the AKM components that we treat as data in our analysis. Section 4 presents the main results in the following order: i) the structure of the social connections; ii) the average causal effects of social connections on hiring together with robustness tests; iii) the effects heterogeneity, by person and establishment effects, and iv) the overall sorting patterns with and without connections. Section 5 presents results from validation analyses showing how connections compete, the role of observable characteristics of the agents, and how matching through social connections is related to future labor market outcomes. Section 6 provides concluding remarks.

2 Empirical strategy

2.1 Set-up, definitions and notation

Because we want to understand the effect of connections on the allocation of jobs, we build our set-up using different components: connections, displacement events, hires. We relate the three using a Rubin-style "causal" model.

Our main ("treatment") variable of interest is a *pair-specific* connection indicator $C_{ijk(l)}$, which equals one if there exists a social connection between worker i, employed in establishment j, and an "intermediary" worker l employed in establishment $k(l)$, and zero otherwise. In the paper, $C_{ijk(l)}$ takes the value one if i and l belong to the same family, have worked together in the past, are former classmates, or are neighbors with children of the same age. To understand the nature of connections, we first describe the distribution of $C_{ijk(l)}$ across combinations of worker and establishment types. Once these networks are described, we estimate the causal effects of social connections on the allocation of workers across establishments. For most of our analysis, we can abstract from the identity of the intermediary worker and focus directly on the worker-to-establishment links using the short-hand notation C_{ijk}.

To assess whether having a connection C to a specific establishment causes workers to move into that specific establishment, we use displacement events due to establishment closures. These events provide us with sets of workers that leave the same firm at the same time under similar (exogenous)

mobility conditions. As discussed in more detail below, using groups of forced movers helps us with identification but it may open questions about how general our results are. We therefore return to a more general case using *all* job-to-job movers in Section 4.4.

Our key outcome is H_{ijk}, which equals one if worker i (from closing establishment j) is hired by establishment k, and zero otherwise.[13] Because of displacements, staying in the closing establishment j is not an option and it is sufficient to concentrate on the probability that establishment k hire worker i (from j) as a function of social connections between k and i.

Using *potential outcomes notation*, we let H^0_{ijk} equal one if i would be hired by k in absence of connection to k. Correspondingly, H^1_{ijk} equals one if i would be hired by k in presence of a connection to k. The *causal effect* of interest is the difference between the two, i.e. $\gamma_{ijk} = H^1_{ijk} - H^0_{ijk}$. Thus, the observed outcome is

$$H_{ijk} = H^0_{ijk} + \gamma_{ijk} C_{ijk}. \tag{1}$$

Both the causal parameter γ and the non-connected hiring outcome H^0_{ijk} capture factors affecting the probability that worker i (from j) is hired by establishment k:

- The causal parameter γ captures all aspects that crucially rely on the existence of a current social connection at the time of displacement. Such aspects may include the displaced worker's knowledge about vacancies in establishment k (if transmitted through connections) as in, e.g., Granovetter (1973) or Calvó-Armengol and Jackson (2004), awareness about job attributes, worker skills and/or preferences that is being transmitted to workers and/or employers through social connections as in, e.g., Montgomery (1991) or Dustmann et al. (2016), or fundamental preferences for working together with a connected worker as in, e.g., Heath (2018).

- The non-connected hiring outcome H^0_{ijk}, on the other hand, encompasses all aspects that are relevant for whether or not displaced worker i is hired by establishment k in the absence of a social connection between the two. This includes all combinations of preferences, job attributes, skills and production technologies which affect the probability that displaced worker i (from j) is hired by employer k even without a social connection between the two.

[13] Formally, we can define an indicator variable M_{ijk} for the event that any worker i moves between j and k and define an indicator S_{ij} equal to one if worker i separates from her previous employer j, and zero otherwise. Then, $M_{ijk} = 1$ if $S_{ij} = 1 \wedge H_{ijk} = 1$ and $M_{ijk} = 0$ if $S_{ij} = 0 \vee H_{ijk} = 0$. Establishment closures ensure that $S_{ij} = 1$ for all workers from j.

2.2 Identification with constant effects

For now, we assume that the impact of social connections is constant (i.e., $\gamma_{ijk} = \gamma$), an assumption we relax in Section 2.3.1 below. Because equation (1) defines connections and hiring outcomes for *pairs* of agents (i.e., worker i and employer k), we adopt a *dyadic* data structure, where each observation is a combination of a displaced worker and an establishment.[14]

To identify the causal parameter γ, we need to account for any systematic relationship between social connections (C_{ijk}) and the potential that connected employers would have hired the very same (connected) workers through the "market", had the connections not been there (i.e. $H^0_{ijk}|C_{ijk} = 1$).[15] Identification through conditioning on observables is feasible if there exists a set of fixed-effects α and a vector of observable factors X_{ik} such that $H^0_{ijk} = \alpha + X_{ik}\beta + \varepsilon_{ijk}$ and $\varepsilon_{ijk} \perp C_{ijk}$, i.e. if the potential for a "market" hire is independent of the presence of a social connection, conditional on fixed-effects and observable factors. For the remainder of the paper, we maintain the $\varepsilon_{ijk} \perp C_{ijk}$ assumption under various fixed effects configurations.

Our main identification strategy relies on a fixed-effect α_{jk} for each *pair of closing establishment (j) and potential hiring establishment (k)*. Using the dyadic data described above, we estimate the following model:

$$H_{ijk} = \alpha_{jk} + X_{ik}\beta + \gamma C_{ijk} + \varepsilon_{ijk}, \tag{2}$$

where the establishment-pair fixed-effects (α_{jk}) account for all shared aspects that relate the closing establishment (j) and the potential hiring establishment (k) to each other. Because an establishment only shuts down once, these fixed-effects are in practise *year-specific*. X_{ik} is a vector of worker characteristics that may make individual i a more suitable hire for establishment k than other workers from j even in the absence of social connections. The model is easily extended to account for different types of connections (e.g. family, coworkers, classmates and neighbors) by replacing γC_{ijk} with $\sum_{n=1}^{N} \gamma^n c^n_{ijk}$ in equation (2), where $c^1_{ik}, \ldots, c^N_{ik}$ are indicators for N types of social relations (family...) measured in the data.

The dyadic set-up of equation (2) resembles Kramarz and Skans (2014), which in turn builds on Kramarz and Thesmar (2013). Following these studies, we estimate the equation using linear probability models and treat each dyad as an independent observation.[16] The strategy to use (j,k) fixed-effects allows us to limit the sample to the sets of observations where there is variation in C_{ijk} within the j,k pair. Thus, we use all possible i,k dyads where establishment k is socially connected to at least one (but

[14] When workers are displaced multiple times, they will form a new set of dyads for each displacement event, as should be evident from what follows.

[15] For simplicity, we refer to such hires (without social connections) as "market hires".

[16] Kramarz and Skans (2014) study parental connections on labor market entry in Sweden. Saygin et al. (2014) uses a very similar set-up to study the importance of former co-workers for labor market outcomes of displaced workers in Austria. Hensvik et al. (2017) uses the setting of Kramarz and Skans (2014) to analyze the role of summer-job connections over the business cycle.

not all) of the workers from i's closing establishment j. Because of the fixed effects, the estimates of interest would not be affected at all if we used the full set of possible combinations of displaced workers and ongoing establishment across the economy instead (but the estimation would not be feasible due to an excessively large sample).

Two aspects of our sample construction are crucial. First, the sample is only constructed based on the pre-existing network-structure and *not influenced by any endogenous outcomes*. Second, the sample-construction includes (essentially) all connected dyads since the number of cases where all displaced are connected to the same establishment is zero for all practical purposes. Thus, the sample-reduction induced by our fixed effects is, in practice, akin to the selection of an appropriate control group (those displaced in the same event), and should *not be interpreted as a restriction that removes any identifying variation of interest*.

Our establishment-pair fixed-effects models allow us to fully and flexibly control for the interaction of all aspects related to location, industry, and firm-specific human capital in each *pair* of employers at each point in time. For this identification strategy to work, we need multiple workers who leave the same establishment at the same point in time under similar conditions. We therefore focus on workers displaced because of an establishment closure. By contrast, had we analyzed the full set of (endogenous) movers in ongoing firms, a sample comprising all movers would be highly selected as 90 percent of the workforce remains in its original workplace every year. Thus, it would *either* have to include a mass of workers who cannot be used for identification because they are highly unwilling to move (the majority, stayers) *or* condition our analysis sample on the endogenous outcome (leaving the original employer). By relying on establishment closures, we ensure that the *decision to leave the original establishment, and thus entering into our sample, is exogenous from every individual worker's perspective*.[17] The cost of using the displacement sample is essentially about potential concerns on how general our results are, i.e. their external validity. We therefore return to the general case in which all movers are included in the statistical analysis in Section 4.4.

An important virtue of the establishment-pair fixed-effects identification is that it can be estimated for a large set of combinations of displaced workers and connected establishments, which is necessary if we want to analyze the relationship to sorting inequality.[18] The identifying assumption underlying this model – i.e. that workers from the same displacement event with and without social connections to a specific employer are equally likely to be hired by that employer through the market, conditional

[17] It is, however, important to note that we are making all our inferences based on comparisons of each individual displaced worker with the group of co-workers displaced during the same event, and we are not comparing to other non-displaced workers as is common in the traditional plant-closures literature.

[18] Since our ultimate goal is to relate the causal effects to sorting inequality, the design of identification strategies using instrumental variables (IV) techniques is extremely challenging. To see why, an IV strategy would require social connections to be randomly allocated to specific establishments across the economy. The most useful source of variation would plausibly rely on mobility of social connections across time, and we show that such identification strategies provide estimates that are very similar to our main estimates for models with homogenous effects.

on X – may be challenged on *a priori* grounds. However, we provide evidence strongly suggesting that our approach provides us with highly reliable estimates. Our confidence in the results relies on three different sets of corroborating exercises:

Specificity: If we replace the actual connected establishment k by other very similar establishment k', the effects shrink toward one-tenth or less of the original effect. We use two different types of such "placebo-style" regressions: *i*) other, randomly selected, establishment within the same location and industry and *ii*) other establishment within the same location and within the same firm. None of these exercises involve a direct social connection to the establishment (but, in case *ii*, to the firm) but should otherwise provide close substitutes in dimensions such as persistent and common preferences, amenities, and skill-match quality.

Robustness to time-invariant individual preferences for specific establishments: We exploit time-variation in workers' social networks between displacement events for the small sample of workers who are displaced multiple times. This variation allow us to estimate a version of the model with α_{ik} (instead of α_{jk}, as in the baseline) fixed-effects, where k is constant over time. These models thus rely on the fact that worker i may have a connection at establishment k during one of her displacement events, but not during another displacement event. Identification with α_{ik} fixed-effects thus accounts for all time-invariant individual preferences for specific establishments.

The role of social proximity: Our data allow us to construct measures that proxy for social proximity, or "tie strength", of a given social relationship. Our estimated effects are consistently larger the larger social proximity between the connected is. This is true regardless of the proxy for tie strength we use. These patterns suggest that social interactions indeed are a crucial ingredient in the process.

2.3 Sorting inequality

To adapt the framework of equation (1) to a study of sorting inequality in the spirit of Card et al. (2013) is straightforward. To do this, we first estimate the AKM model of Abowd et al. (1999):

$$\ln w_{it} = \theta_i + \psi_{k(i,t)} + X_{it}\beta + \varepsilon_{it}, \tag{3}$$

where w_{it} is worker i's wage in year t, θ_i is a person fixed-effect for worker i, and $\psi_{k(i,t)}$ is an establishment (k) fixed-effect in year t. $X_{it}\beta$ is a vector of control variables, which as in Card et al. (2013), includes an unrestricted set of year indicators, and education level interacted with age in quadratic and cubic terms.[19]

The estimates of θ_i and $\psi_{k(i,t)}$ will be treated as data, following Card et al. (2013) among others. To ensure that the impact that social connections may have on post-hiring wages is not transmitted

[19]Education level is categorized in three levels: compulsory school, high school, and college/university. Age is normalized relative to age 40.

into the estimates of the person effects of the displaced workers, we estimate these effects on data for the years *preceding* job displacement (i.e., separately for each displacement-year cohort). This also ensures that errors in the estimated person effects of the displaced are unrelated to estimation errors in the establishment effects of potential future employers, as these are estimated in different samples. Thus, the mechanical negative association between estimation errors in simultaneously estimated person and establishment effects discussed by, e.g., Andrews et al. (2008) will not impact our estimates.[20] To further ensure the validity of our strategy and the resulting estimates, we also present evidence based on a) estimates where employers are grouped by the clustering algorithm of Bonhomme et al. (2018), b) estimates where employers are grouped by the (firm-level) value-added per worker.

To understand the role of social connections for the unequal access of workers to jobs ("sorting inequality"), we start by analyzing the distribution of connections between the various types of agents. We correlate the person-effect of displaced workers (i) with the person-effect of their social connections (i.e., the intermediary workers l).[21] A positive correlation, i.e., $\text{Corr}(\theta_i, \theta_l \mid C_{il} = 1) > 0$, will be interpreted as *baseline homophily*. We also analyze the correlation between the displaced workers' person effects and the types of establishments they are connected to, i.e., $\text{Corr}(\theta_i, \psi_k \mid C_{ik} = 1)$. A positive correlation implies that the structure of social networks contribute to sorting inequality.

Then, we estimate how the causal effects of connections relate to the person and establishment effects. To do so, we let $\gamma_{ijk} = \gamma(\theta, \psi)$ in equation (1) and identify these effects through an extension (described next) of the strategy outlined in equation (2).

2.3.1 Identification when causal effects of connections vary with the AKM components

We are interested in how the effects of connections vary with the estimated AKM person and establishment effects of the connected agents.[22] Thus, we extend equation (2) to yield:

$$H_{ijk} = \alpha_{jk} + X_{ik}\beta + \gamma(\theta_i, \psi_k)C_{ijk} + \varepsilon_{ijk}, \tag{4}$$

where we assume that $\gamma(\theta_i, \psi_k)$ is a second-order polynomial:

$$\gamma(\theta_i, \psi_k) = \gamma_1\theta_i + \gamma_2\theta_i^2 + \gamma_3\theta_i\psi_k + \gamma_4\psi_k + \gamma_5\psi_k^2. \tag{5}$$

[20]This is potentially important because our measures of sorting concerns the relationship between person effects and the establishment effects of new potential employers. In the end, we do, however, show that the sample splitting is in fact not crucial for our qualitative conclusions. Kline et al. (2018), show how variances and covariances can be corrected for estimation errors using cross-sectional split-sample estimation but we are not employing these methods as our purpose is to extract the components, and not to compute the aggregate statistics.

[21]Essential here to remember: the two sets of effects come from different samples and are therefore not affected by the so-called limited mobility bias.

[22]Extensions to other sources of heterogeneity discussed in the paper are straightforward.

Three important remarks about this set-up must be made:

1. The estimates of γ_3 identifies the contribution of causal effects to sorting inequality since they govern the role of interactions between person and establishment effects.

2. The control vector X_{ik} includes θ_i, θ_i^2 and $\theta_i\psi_k$ but *neither* ψ_k *nor* ψ_k^2 as these two last terms are captured by the establishment-pair fixed-effect α_{jk}.

3. In specifications where we focus on a single quality dimension (i.e., either person *or* establishment effects), we categorize the person/establishment effects through ten unrestricted decile indicators, instead of the more restrictive polynomial representation.

2.3.2 Within-worker identification

Displaced workers tend to be endowed with multiple social connections, tying her to different establishments. We can use this fact to set up an alternative model that directly assesses the relative importance of different connections *for this particular worker during a specific displacement event*. This "within-worker" perspective is new to the literature, but well-suited for our analysis of how the impact of connections vary with aspects of the agents or the type of connection.

Let us first define K_{ijk} as the set of establishments k that displaced worker i (from j), is connected to (i.e., define $K_{ijk} : k \in K_{ijk}$ iff $C_{ijk} = 1$). To identify the relative importance of the various types of connections or agents within this set, we define a fixed-effect $\phi_{K_{ijk}}$ for each K_{ijk} set.[23]

Thus, focusing on heterogeneity in terms of agents (not connections) for reasons of brevity, we can rewrite equation (4) as:

$$H_{ijk} = \phi_{K_{ijk}} + X_{ik}\beta + \gamma^I(\theta_i, \psi_k)C_{ijk} + \varepsilon_{ijk}, \tag{6}$$

where, as in (4), γ^I represents a second-order polynomial and X_{ik} contain the components of the same polynomial.[24] Because the new fixed effect requires the existence of a social connection between worker i and establishment k, some components of these polynomials are in practice absorbed.[25] Thus, equation (5) reduces to:

$$\gamma^I(\theta_i, \psi_k) = \gamma_3^I\theta_i\psi_k + \gamma_4^I\psi_k + \gamma_5^I\psi_k^2, \tag{7}$$

[23] As before, identification only comes from the observations (dyads) for which there is variation in c_{ik}^n within the fixed-effect set K_{ijk}. Conceptually, there exists a corresponding set of individual fixed-effects for dyads without connections but within this set there is by definition no variation in c_{ik}^n.

[24] Comparing $\phi_{K_{ijk}}$ to α_{jk}, we note that i is a subset of j, and all k's that are included in K_{ijk} are covered by some α_{jk}.

[25] Since the estimates are derived conditional on the individual's (fixed) set of connections, the model will not capture the direct impact of the worker's quality either (i.e., the person effect cannot be included in X_{ik}).

where the parameters are numbered as in equation (5) and the estimate of $\gamma_3^{n,I}$ thus identifies the contribution of causal effects to sorting.

As with equation (2) above, we can introduce a sequence of different types of connections (family, coworkers,..), each indicated by c_{ijk}^n instead if C_{ijk} but due to the $\phi_{K_{ijk}}$ fixed-effects we need to leave out one *reference-type* of connections when estimating this model.

Equation (6) allows us to examine the interaction of the AKM effects of the connected establishment or the type of connection *for a given worker* at a specific displacement event. Identification comes from workers with connections to both high- and low-wage establishments, or with different types of connections. This empirical strategy allows us to handle the fact that the types of workers who are connected to high-wage establishments may be different from those who are connected to low-wage establishments, and that different types of workers have different types of connections.[26]

3 Data and definitions

3.1 The administrative registers

The analyses are based on administrative data for the entire Swedish population covering 1985–2009. Administrative records are linked using anonymized identifiers at the individual-, establishment-, and firm-levels. The main data source is an employment register (*Registerbaserad arbetsmarknadsstatistik*) with information from the National Tax authority. The statutory income statements, filed to the taxation authorities by the employers, identify both the employee and the establishment's organization, which allows us to link all employees to their employer. The social connections are identified using information on family trees from population-wide birth records (*Flergenerationsregistret*), information on household members and neighbors from Statistics Sweden's longitudinal database (*LOUISE*), and information on graduation classes from high school and college/university from graduation registers (*Skolregistret* and *Universitets- och högskoleregistret*).

3.2 Defining closing establishments and displaced workers

We identify establishment closures in the data in two steps: First, we select all establishments with a non-missing identifier in November of year t, but whose identifier is no longer in the data in November year $t+1$. Second, we eliminate cases where the establishment identifier is deemed to be missing for other reasons than that the establishment ceased to operate (e.g., mergers and dispersals).[27] We retain

[26] See see Section 3.3.1 below for a detailed discussion about the coverage.

[27] We follow Hethey-Maier and Schmieder (2013) and define "true" closures (or "atomized deaths") as those where no cluster of more than 30 percent of the workforce at the exiting establishment in year t are found at the same establishment in year $t+1$.

only the closures of single-establishment firms that occurred in the private sector during 1990–2006 and involved at least four employees. [28] The displaced workers are consequently defined as those, of ages 20–64 years, who in November of year t are employed at an establishment that closes down during the following 12 months. This procedure results in 31,538 closing establishments j and somewhat less than 289,332 displaced workers i.

3.3 Defining social connections

We consider four broad types of social connections: family members, former co-workers, former classmates, and current neighbors. Below we give the precise content for each of these groups. Throughout, we restrict the data to those cases where we can be reasonably sure that the measured connections are valid, i.e., that the people involved actually did meet each other. In particular, in cases where groups of agents are very large, we prefer to not code them as social connections in order to not risk including false connections into our data. Details are spelled out below.

Family members include parents, adult children, spouses, and siblings (either full or half). We have relied on birth records (which are near complete for the Swedish born) to identify parents, children, and siblings. Spouses are defined by household indicators, which capture those who resided together and who either were married or had children in common.

Former co-workers comprise workers who were employed at the same workplace *before the current one i.e., not the closing establishment*. We limit former co-workers to those in the most recent of past workplaces, using data going back to 1985.[29] If a workplace is very large, the measured connections are likely to be very imprecise and noisy. Therefore, we constrain the data to cases with less than 100 employees at the former workplace.[30]

Former classmates are identified at high-school and/or at college/university. High-school students are tracked into different occupational programs that usually are offered as one class per school and program combination. Therefore, we identify former classmates from high-school as those who shared school, program, and graduation year. Students from university are similarly identified as those who graduated at the same college/university, within the same field/major, and during the same year. Because the graduation records are not available prior to 1985, we have information on former classmates only

[28] In practice, we exclude the public sector defined as all organizations with 2-digit institutional codes of 11–14 or 3-digit institutional codes 151, 152, 501 and 502 before 1999, and all firms with 1-digit institutional codes of 3–5 or 3-digit institutional code 721 thereafter. The narrower sampling period, rather than the full 1985–2009-period, allows both for a pre-closure period when social connections are formed and for a potential post-hire period. Counted as employees are those for whom the particular establishment is their main workplace (i.e., the establishment in November from which they receive the largest annual earnings).

[29] We only consider each employee's main workplace (i.e., establishment) in the month of November.

[30] This exclude 25 percent of the former co-workers. The size requirement, which we also impose on former classmates and current neighbors, is also needed for computational reasons. Otherwise, the number of dyads between all displaced workers and potential hiring establishments would explode.

for the younger cohorts. However, our results indicate that the value of former classmates depreciates fairly rapidly over time (see Section 5.2). There are also cases where we have failed to identify what could reasonably be defined as a class, because when a school catered large cohorts within one field, it was presumably divided into two (or more) classes, which cannot be observed in the data. To reduce the influence of pure measurement error in our measured connections we, therefore, remove the cases where more than 100 former students are found within the same (constructed) class (i.e., analogous to the procedure for former co-workers).[31]

Current neighbors are defined as those residing in the same area according to Statistics Sweden's neighborhood indicator SAMS (Small Areas for Market Statistics). There are about 9,200 such areas in Sweden, and each contain, on average, approximately 1,000 residents. Hence, the identified networks of neighbors would in most cases extend far beyond the group of people who actually interact with each other. In order to define measures of residential networks that capture social interactions, we have therefore included only those who both reside in the same SAMS area *and* have children in the same age group.[32] The intended logic is that parents with children in the same age group are more likely to meet (or have met) at playgrounds, schools, or other local child activities. This notion receives strong empirical support in Bayer et al. (2008) that shows that neighbors with same-aged children are substantially more likely to work together than other neighbors. Analogously to former co-workers and classmates the groups of current neighbors with children of the same age are constrained to those containing less than 100 people to reduce the impact of measurement errors.

Three additional requirements have been imposed on all social connections defined above: each individual connected to a displaced worker must be (i) 20–64 years old, and (ii) employed at a private sector establishment (with an associated identifier in the data) in November of both years t and $t+1$.[33] All restrictions were imposed after applying the group size constraints of 100 that were described above.

Using these definitions of social connections, we connect each displaced worker to, on average, 8.8 potential hiring establishments k, whereof 1.2 through family members, 2.6 through former co-workers, 3.5 through former classmates, and 1.6 through current neighbors (see Table 1).[34]

[31] This excluded 21 and 10 percent of high school and college/university "classmates", respectively.

[32] Using Statistics Sweden's child age groups: 0–3, 4–6, 7–10, 11–15, 16–17, or 18+ years.

[33] The strategy ensures that the connected intermediary worker was employed at the particular establishment at the time of the (potential) hire of the displaced worker. Workers who are employed without a well-specified geographic location (e.g., home care workers) lack the establishment identifier.

[34] Recall that we only count connections that are of working age, be employed at the same private sector establishment in t and $+1$.

Table 1: The number of potentially hiring establishments (k) (per displaced worker (i) and in total) connected to displaced worker (i), by type of social connection

	Type of social connection				
	Family member	Former co-worker	Former classmate	Current neighbor	Any connection
Per displaced worker (i)	1.17	2.59	3.52	1.63	8.82
In total	338,518	749,370	1,019,027	471,032	2,553,066

Note: The Table gives the number of connections for displaced workers within the used sample (1990-2006).

3.3.1 Incomplete coverage of social networks and measures of tie strength

We do not observe the full set of actual social connections. This limitation is shared with all other studies using data from administrative registers, or social media platforms, as measures of social networks since none of these sources record the entirety of social relations. However, under the key assumption that unrecorded friendships have the same properties as the measured types of connections, this should not affect our conclusions. The reason is that we are not measuring the aggregate effects of network size. Note that:

- When we study the sorting patterns among connections (searching for "baseline homophily"), we correlate person and establishment effects among the measured connections, thus estimating $\text{Corr}(\theta_i, \psi_k \mid C_{ik} = 1)$. This analysis has general validity as long as our observed connections have the same properties as unrecorded friendships. The analysis would, however, be distorted if we included false connections into the analysis.

- Causal effects are identified by comparing observed connections to dyads without observed connections using the equation $H_{ijk} = \alpha_{jk} + X_{ik}\beta + \gamma C_{ijk} + \varepsilon_{ijk}$. The dyad ($\alpha_{jk}$) fixed-effects identification ensures that estimates can only be biased through random measurement errors if many other workers have unrecorded connections within the set of dyads we analyze, i.e. to an employer where one (or more) co-displaced already has an observed connection. Such unrecorded connections would erroneously attribute some connected dyads to the set of "market matches" (H^0_{ijk}). Thus, the market match frequency would be overestimated and the causal effects γ_{ijk} underestimated. Since the estimated baseline ("market") probability of entering firms to which others have connections empirically turn out to be very low, this does not appear to be an important concern.

- However, when studying overall sorting patterns in Section 4.4, incomplete coverage should attenuate our estimates since our residual category of "market matches" will contain a mixture of true market matches and unrecorded social connections. We discuss this concern in Section 4.4.

We cannot observe family members, former co-workers, former classmates, and current neighbors

for all workers within our sample since some workers do not possess all types of connections (e.g. some workers do not have a spouse, some workers were never employed before and therefore do not have any former co-workers, not all workers went to university, and so forth). Furthermore, our data are incomplete since we lack graduation records for the oldest cohorts of workers and because current neighbors with children in the same age can only be identified for workers who actually have children. And, by our restriction, connections formed at very large establishments, neighborhoods or classes are removed because only a few of the agents within them will actually have interacted. However, our *within-worker identification strategy*, allows us to directly address this issue. The strategy precisely compares the role of different types of connections for workers that have multiple types and thus removes all the impact of differences across workers with different types of connections.

Part of our analysis uses the concept of tie strength, building on Granovetter (1973)'s observation that social relations differ in the intensity of their mutual interactions, intensity which matters for the usefulness of connections in job-search. Pairs of agents with frequent mutual interactions are typically labelled as having "strong ties". Pairs of agents who interact less are instead labelled as having "weak ties". As in Gee et al. (2017), we expect each connection to matter more if there is more interaction (i.e. if the tie is stronger), even though the weaker ties can have the advantage of providing more information that is previously unknown to the job seeker, as emphasized by Granovetter (1973). In this paper, we measure connections in several domains (i.e. family, former co-workers, former classmates, current neighbors) to get as broad a coverage as possible, but within each of these domains we consciously seek to restrict the data to those cases where it is plausible that agents have some minimal level of interaction. We therefore only use neighbors with children of the same age and restrict the sample to "contact-forming" groups with less than 100 persons. Although we cannot measure the frequency of interactions directly, we hypothesize that connections that are relatively recent and/or formed in durable and small groups are associated with more interactions (i.e. ties are stronger). We show results (see in particular Section 5) suggesting that the frequency of interaction indeed is a crucial determinant of the quantitative importance of connections. The results suggest that by removing connections formed in very large groups, we mostly remove connections that are too weak to be useful in the job-search process.

3.4 Earnings data and AKM estimates

When estimating the AKM model we first use the universe of workers and establishments for our full data period 1985-2009. The sample and estimation results are described in Table 2.[35] The table has five columns. In the first we show estimates for the full sample. In the second, we focus on all job-

[35] See Appendix Figure A.1 for a 3d graph of the joint distribution of person and establishment effects deciles.

to-job transitions during our main sample period (1995-2006) and display the person effects and the establishment effects for the sample of private sector hiring establishments. As discussed in Section 2.3, we use pre-hire data to estimate these person effects to avoid the impact of both reverse causality (from connected hiring to wages) and the simultaneity bias arising from correlated estimation errors.[36] We discuss estimates from alternative classification strategies relying on Bonhomme et al. (2018) and valued added per worker in the results section. In the third, we focus on job-to-job transitions that are associated with social connections as measured in our data. Fourth, we focus on the sample of new hires after displacements. And in the fifth, we focus on connected hires after displacements. These samples are somewhat smaller than the overall displacement samples since we need the workers to be employed the year before displacement (or earlier) in order to estimate the person effects.

Table 2: Description of the AKM samples

				Estimation sample	
	AKM estimation sample	All job-to-job hires	Connected job-to-job hires	Displaced hires	Displaced connected hires
Number of person effects	5,785,081	3,315,423	424,789	83,537	10,202
Number of establishment effects	829,111	276,298	119,377	43,469	7,127
Mean of person effects	0.000	-0.104	-0.101	-0.117	-0.111
Mean of establishment effects	4.502	4.525	4.525	4.507	4.503
Std dev. of person effects	0.270	0.269	0.255	0.250	0.245
Std dev. of establishment effects	0.126	0.154	0.152	0.155	0.152
Correlation person/establishment effect	0.038	0.129	0.083	0.089	0.049
Mean of log wages	9.664	–	–	–	–
Std dev. of log wages	0.466	–	–	–	–
No of observations	62,002,038	3,315,521	424,792	83,540	10,202

Note: Column (1) presents summary statistics for AKM person and establishment effects in the full sample of workers and establishments during 1985–2009. In columns (2)–(5) we present statistics for our observation period of interest 1995–2006. When focusing on transitions in columns (2)–(5), we use the AKM person effects estimated in the pre-transition period. Displaced workers are defined in Section 3.2. Connected hires include the four broad types of social connections (family members, former co-workers, former classmates and current neighbors) defined in Section 3.3.

Table 2 shows that the person effects of recent hires are lower than average and that the dispersion of person effects is lower in the samples that were hired through connections, both in general (all hires) and after displacements. Furthermore, estimated correlations between person and establishment effects suggest that connected hires are less sorted than the overall samples of connections, both in general (all hires) and after displacements. We return to this issue in Section 4.4.

3.5 Estimation data (dyads)

To create our estimation data, we form pairs between each displaced worker i (from closing establishment j) and each potential hiring establishment k. We include all pairs where *someone* from estab-

[36] Thus, estimates are drawn from year-specific regressions on samples that go from the first year of the sample period to the year just before the displacement.

lishment j has a social connection to establishment k. The sample construction thus only rely on the pre-displacement network structure and is unrelated to endogenous outcomes and we do not remove any of the identifying variation (except for a trivial number of cases where all displaced form an event are connected to the same establishment). The procedure generates a data set comprising 41 millions pairs of workers and establishments. Within these data, some workers from establishment j will have an actual connection to establishment k whereas others will not, allowing us to estimate the effects of connections within a model with j,k-pair fixed-effects.

Table 3: Summary statistics for the displaced and intermediary workers and for the estimation sample comprised by dyads of potential hiring establishments k and displaced workers i

	Displaced workers (i)		Intermediary workers (l)		Dyads of displaced workers i and potential hiring establishments k	
	N	%	N	%	N	%
Sex						
Female	113,738	39.31	853,104	37.87	24,119,673	58.67
Male	175,594	60.69	1,399,683	62.13	16,992,101	41.33
Nativity						
Swedish born	257,686	89.06	2,095,419	93.01	37,292,005	90.71
Foreign born	31,646	10.94	157,548	6.99	3,819,769	9.29
Age						
20–34 years	149,529	51.68	1,379,492	61.23	25,565,797	62.19
35–49 years	88,075	30.44	629,184	27.93	10,461,203	25.45
50–64 years	51,728	17.88	244,291	10.84	5,084,770	12.37
Attained education						
Compulsory school	69,044	23.86	312,826	13.89	6,538,236	15.90
High school	163,735	56.59	1,374,302	61.00	24,250,767	58.99
College/university	54,829	18.95	561,351	24.92	10,199,868	24.81
Employment in $t+1$						
Any employment	211,282	73.02	2,252,967	100.00	33,921,201	82.51
In connected establishment	9,143	3.16	2,252,967	100.00	17,707	0.04
AKM person effects [a]						
Low	78,098	26.99	1,003,740	44.55	11,711,626	28.49
Medium	65,505	22.64	733,727	32.57	8,495,645	20.66
High	46,444	16.05	445,638	19.78	4,756,028	11.57
N/A	99,285	34.32	69,862	3.10	16,148,479	39.28
AKM establishment effects [a]						
Low	102,714	35.50	509,163	22.60		
Medium	88,522	30.60	795,413	35.31		
High	74,460	25.74	916,296	40.67		
N/A	23,636	8.17	32,095	1.42		
No of observations	289,332		2,252,967		41,111,774	

Note: The Table gives descriptive statistics of displaced workers (i), intermediary workers (l) and dyads of displaced workers and potential hiring establishments (i,k) within the used sample (1990-2006). The dyads make up the used sample for the causal analysis.
[a] See Section 2.3 for details on the AKM model

Data are described in Table 3 with separate columns describing the 289,000 displaced workers, the 2.2 million intermediary workers, and the 41 million pairs of displaced workers and potential hiriong

establishments.[37] Because we focus on private-sector closures, two-thirds of displaced workers are male. Furthermore, we find that 90 percent are Swedish born and half of the workers are below age 35, and 74 percent have at least a high school degree. 73 percent find employment within the year after displacement.

Since we require a pre-period to estimate the AKM effects of the displaced, we do not have such estimates for about a third of all displaced workers. In Section 4.4 we discuss results where we also include "post rehire" data in the estimation of the AKM-effects and results are in fact very similar.

A much larger share of intermediary workers are employed in high-wage establishments relative to the displaced workers. Notably, the 41 million pairs in our data only make up a small subset of all possible combinations of displaced workers and (non-connected) potential hiring establishments, and matches were obviously formed between displaced workers and establishments not included in this data set. However, given the establishment(-pair) fixed-effects included in our models, hiring establishments without any connection to a closing establishment will not contribute to the identification of the parameter of interest, hence are not included.

Our data set contains nearly 300,000 displaced workers and over 300,000 plants per year, a complete set of dyads would therefore comprise around 90 billion observations. Our data do, however, cover a large part of the economy. Each year it contains about 50,000 unique establishments, and for the full data period there are 289,000 unique establishments. By construction, left-out establishments tend to be very small in terms of employment.

4 Results

In this Section, we present the results from analyses of (displaced) workers' social connections and how they affect the matching to the potential hiring establishments. We present the results in the following order:

1. We describe displaced workers' social connections in Subsection 4.1,
2. We estimate the impact of connections, including robustness tests supporting a causal interpretation, in Subsection 4.2,
3. We interact the causal effects with person and establishment effects in Subsection 4.3, and
4. We provide an overall assessment of the role of connections in sorting inequality, both for the displaced and for all hires, in Subsection 4.4.

[37] Further statistics at the establishment level are provided in Appendix Table A.1.

4.1 Sorting of social connections

We now document the sorting patterns of connections between establishments and displaced workers. We use the data on all connections we observe between displaced workers and ongoing (potential hiring) establishments. To describe the underlying sorting patterns, we first calculate correlations between the person effects of displaced workers (θ_i) and the person effects of the intermediary workers to whom they are socially connected (θ_l) and then show the corresponding correlation between θ_i and the establishment effect (ψ_k) of l's employer. Thus, for the sample of displaced workers, we show how their "quality" (in terms of person effects) correlates with both the "quality" of their social connections and the "quality" of these workers' employers.

Table 4: Correlations between the AKM person effect of the displaced worker and the person and establishment of their connections using raw (non-residualized) and residualized person effects

	$\text{Corr}(\theta_i, \theta_l \mid C_{il} = 1)$		$\text{Corr}(\theta_i, \psi_k \mid C_{ik(l)} = 1)$		
	Raw person effects	Residualized person effects	Raw person effects	Residualized person effects	N
Panel A:					
Any connection	0.164	0.071	0.061	0.008	1,597,994
Panel B:					
Family member	0.042	0.057	0.027	0.017	195,099
Former co-worker	0.185	0.098	0.097	0.062	501,706
Former classmate	0.201	0.055	0.084	-0.030	587,511
Current neighbor	0.073	0.053	0.026	0.018	304,678
Panel C:					
Family member					
Parent	0.052	0.031	-0.017	-0.013	37,477
Adult child	0.101	0.046	0.086	0.066	27,265
Spouse	-0.134	0.050	0.032	0.029	25,614
Sibling	0.159	0.072	0.019	0.009	104,743
Former classmate					
High school	0.099	0.047	0.024	-0.033	509,471
College/university	0.289	0.098	0.137	0.013	78,040

Note: Column (1) shows the correlation between the AKM person effect of displaced worker (i) and the person effects of the intermediary workers l to whom they are sociall connected. In column (2) we show the relationship when both these person effects have been residualized from age at displacement, education level and gender. Column (3) shows the correlation between the AKM person effect of displaced worker i and the AKM establishment effects of connected establishments k. In column (4) we have residualized the AKM person effect of the displaced worker (i) from age at displacement, education level and gender.

Results are presented in Table 4. For the average connection in our data, the correlation between person effects of the displaced and intermediary workers is positive at 0.164. Thus, the network structure exhibits positive "baseline homophily" in terms of earnings capacity; "good workers" know other "good workers" as presumed in many standard network models (e.g., Montgomery, 1991). The correlation between person effects of the displaced worker and the establishment effects of the connected employer is also positive, but notably weaker (0.061). These magnitudes can be compared to the correlation between displaced workers' person effects and the establishment effects of their new employers (thus

calculated for the rehired portion of all displaced) which is 0.083 as shown in Table 2 above. Thus, the average within-network sorting between displaced workers and their connected employer (0.061) is nearly as large as the average sorting between workers and re-hiring employers (0.083) for (the rehired portion of) the same sample. By the same metric, the average sorting between persons within networks (0.164) is twice as large as the sorting between person effects and the establishment effects of the new employer.

Panel B shows separate correlations for family members, current neighbors, former co-workers and former classmates. Again, all person-to-person correlations are positive, suggesting that homophily is a general phenomenon, i.e., high-wage workers are connected to other high-wage workers for all our observed connections at this level of aggregation. For all these measures of connections, displaced workers with higher (pre-displacement) person effects are also connected to higher-wage employers. The correlations are clearly largest for the "professional" ties formed at school or in workplaces. The more "socially"-oriented family and neighborhood ties are less sorted.

In Panel C we disaggregate the types of connections even further and here it is obvious that demographic patterns are strongly associated with some specific types of connections; in the data spouses of men are women and vice versa, and gender wage disparities are clearly reflected in the highly *negative* spousal sorting. To assess the importance of these issues, the table shows separate columns for results based on "residualized" person effects where the impact of age at displacement, education and gender have been removed as in Abowd et al. (1999). After this "residualizing" of the person effects, the correlations for spouses turn positive, reflecting positive assortative mating on the marriage market. Panel C also shows that much of the positive sorting for former classmates are driven by by very strong homophily within the network of former university classmates (0.289); in contrast, the former high school classmates are much less sorted (0.099).

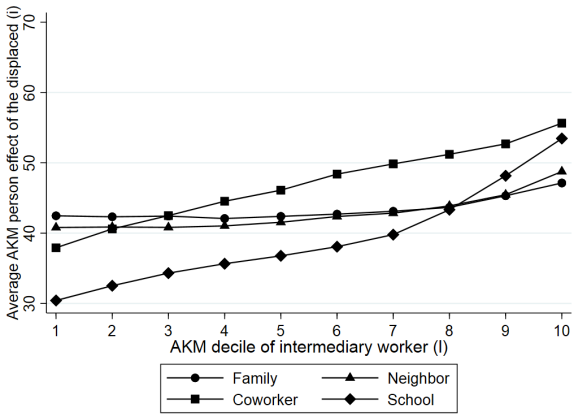
To make this description parallel to the analysis of causal effects presented later, we also perform a less parametric analysis of the social networks from an employer-side perspective. We proceed as follows. We calculate the average person effect of the displaced separately by *i)* decile of the intermediary workers' person effects and *ii)* decile of the connected employers' establishment effect. Results are presented in Figure 1, where each panel corresponds to one column of Table 4. The relationships mostly appear to be linear, at least after residualizing the data, but there is a tendency for accelerated sorting at the top for the school and co-worker connections when using the non-residualized data.[38]

Overall, this analysis clearly shows that high-wage workers are more likely to be connected to other high-wage workers, in particular when these connections are co-workers or classmates (from university). This baseline homophily within the social network also imply that high-wage workers are more likely

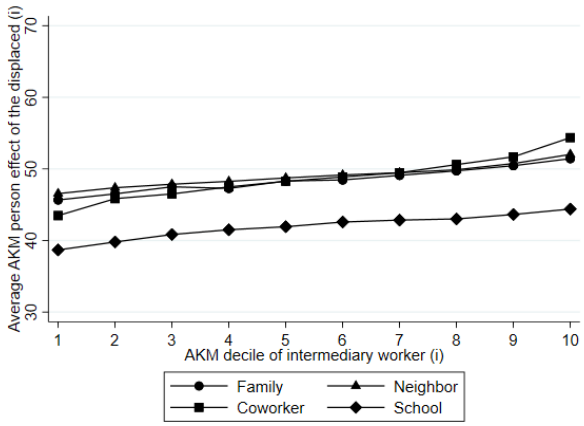
[38]Results for more detailed types of connections (parents, spouses,...) are found in Appendix Figures B.1 and B.2.

to be connected to high-wage establishments. The results thus imply that high-wage workers' social networks comprise better employers than low-wage workers' networks do.

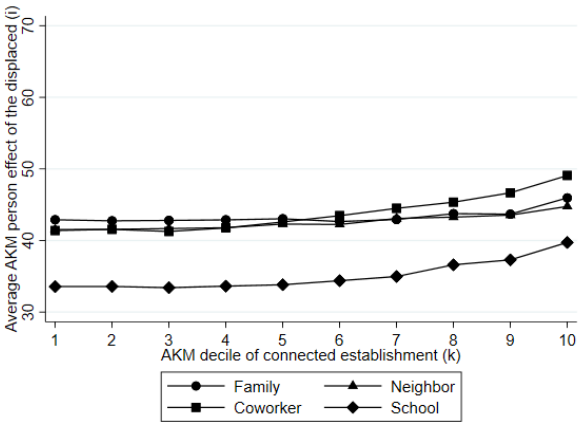
(a) Non-residualized AKM effects: Displaced workers by intermediary workers (col. [1] of Table 4, Panel B)



(b) Residualized AKM effects: Displaced workers by intermediary workers (col. [2] of Table 4, Panel B)



(c) Non-residualized AKM effects: Displaced workers by potential hiring establishments (col. [3] of Table 4, Panel B)



(d) Residualized AKM effects: Displaced workers by potential hiring establishments (col. [4] of Table 4, Panel B)

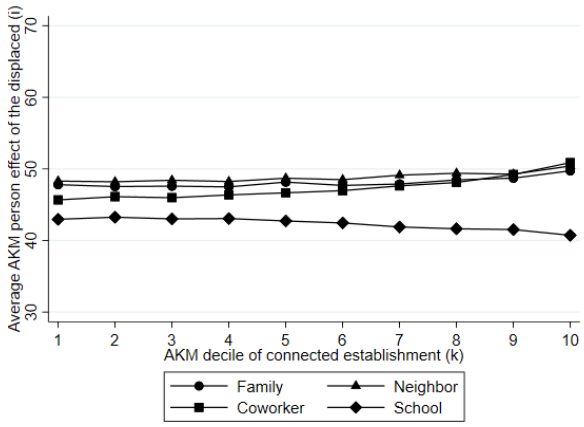


Figure 1: Average percentile of displaced workers' non-residualized (left hand side) and residualized (right hand side) AKM effects (θ_i and $\hat{\theta}_i$, respectively) by decile of intermediary workers' and the potential hiring establishments' AKM effects ($\theta_l \mid C_{il} = 1$ and $\hat{\theta}_l \mid C_{il} = 1$, respectively) for each type of social connection.

4.2 Causal impact of connections on hiring

Table 5 (left column) reports our main estimates of the average impact of social connections (i.e., γ in equation 2) for different subsets of connections. The baseline result suggests that displaced workers are 0.27 percentage points more likely to be hired by each connected establishment k relative to other displaced workers from the same closing establishment j. This effect may appear small, but recall that we are predicting the very rare event that a given worker is entering into a specific establishment as a function of a connection in that very establishment.[39] Standard errors are clustered at the level of the

[39] See estimates of individual neighbors in Bayer et al. (2008), or single moved coworkers in Hensvik et al. (2017), which are of similar magnitudes.

potential hiring establishment. We have verified that the key conclusions remain valid if we instead cluster on both the potential hiring and the closing establishment.[40] To put the estimates in perspective, the estimated effect (of an average connection) is 10 times the baseline probability of hiring by the non-connected (i.e., the constant) of 0.026.

Table 5: Estimated effects of social connections by type of connection, with alternative fixed-effects

	"Baseline" (jk-fixed-effects)		"Multiple displacements" (ik-fixed-effects)		"Within-worker identification" (K_{ijk}-fixed-effects)	
	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)
Panel A:						
Any connection	0.270	(0.005)	0.305	(0.025)	N/A	
Constant	0.026	(0.000)	0.085	(0.012)		
Panel B:						
Family member	1.095	(0.020)	1.196	(0.173)	1.464	(0.036)
Former co-worker	0.253	(0.010)	0.477	(0.048)	0.714	(0.030)
Former classmate	0.066	(0.004)	0.101	(0.025)	0.157	(0.026)
Current neighbor	0.086	(0.008)	0.040	(0.033)		
Constant	0.027	(0.000)	0.075	(0.013)	0.014	(0.002)
Panel C:						
Family members						
Parent	1.867	(0.052)	1.692	(0.440)	2.061	(0.058)
Adult child	0.667	(0.051)	0.797	(0.357)	1.388	(0.065)
Spouse	1.971	(0.078)	2.683	(0.741)	2.526	(0.081)
Sibling	0.696	(0.023)	0.832	(0.195)	1.105	(0.042)
Former co-worker						
1–20	0.527	(0.029)	0.960	(0.124)	1.164	(0.045)
21–50	0.217	(0.018)	0.378	(0.073)	0.664	(0.040)
51–100	0.140	(0.014)	0.224	(0.060)	0.545	(0.038)
Former classmate						
High school	0.064	(0.004)	0.094	(0.025)	0.231	(0.027)
College/university	0.088	(0.018)	0.195	(0.117)	0.464	(0.043)
Current neighbor						
1–20	0.072	(0.072)	0.154	(0.428)		
21–50	0.084	(0.021)	0.025	(0.087)		
51–100	0.079	(0.009)	0.037	(0.035)		
Constant	0.026	(0.000)	0.073	(0.013)	0.010	(0.002)
No of fixed-effects		2,087,560		147,996		548,820
No of observations		41,111,774		295,992		41,111,774

Note: Data are in dyadic form with one observation per combination of displaced worker and potential hiring establishment. All estimates are expressed in percentage points, i.e., the coefficients are multiplied by 100. Baseline model uses a fixed-effect for each pair of closing and potential hiring establishments. Since establishments only close once, these fixed-effects are year-specific by construction. Alternative model for workers with multiple displacements uses fixed-effect for each combination of individual and connected establishment, thus exploiting variation in network structure between displacements. Since establishments only close once, these fixed-effects are year-specific by construction. Within-worker identification uses fixed-effect for each combination of individual, set of connected establishments, and displacement year as discussed in Section 2.3.1 Standard errors are clustered on the potential hiring establishment-and-year level.

The estimated average effect of a connection masks considerable heterogeneity across types of connections. The impact of family members is the largest; a family member within an establishment raises the hiring probability by, on average, 1.1 percentage points (Panel B). The effect varies by type of family

[40] These two-way cluster estimates are available on request.

member from nearly two percentage points for parents and spouses to just over half a percentage point for adult children and siblings (Panel C). Former co-workers have the second largest effects. Having a former co-worker within an establishment increases the hiring probability into that establishment by 0.25 percentage points. Even though former classmates or current neighbors also matter, the magnitudes of the effects are substantially lower (0.07 and 0.09 percentage points).

The differences in results between types of connections potentially reflect the frequency of interactions within each network, i.e., tie strength in the terminology of Granovetter (1973), as family members can be presumed to interact more frequently than, e.g., people who went to school together. In the extreme, where people do not interact at all, the connections should be useless by default. To more directly test this presumption, we separate the co-worker connections by size of the establishment where the connection was formed in Panel C. The results show that connections matter more if formed in establishments with less than 20 employees than if formed within larger establishments. We expand on this analysis even further in Section 5.2 where we show that group size matters even *within* the 1–20 group (a twice larger effect for 1–10 than for 11–20) where one can assume that everyone knew each other while they were working together. Hence, the results support that frequency of interaction is crucial.[41]

The results just discussed are based on a model with (j,k) establishment-pair fixed-effects (left column of Table 5). Next, we discuss a sequence of robustness tests.

Robustness to time-invariant preferences for specific establishments. The middle column of Table 5 shows estimates from a model estimated on the small sample of workers who are displaced multiple times. We restrict the sample to those displaced exactly twice. For this sample, we can exploit changes in a worker's network structure of connections between displacements. We introduce a fixed effect for each combination of displaced person and establishment instead of the baseline j,k fixed effects. The regressions include all dyads where the establishment existed during both displacements to which the worker was connected at least once. We control for the size of the establishment at the time of displacement. Since these models include fixed effects for combinations of displaced workers and connected establishments, the estimates are robust to any time-invariant individual preferences for specific establishments. The results are slightly less precise than in the baseline, but the estimates are essentially unchanged; the effect of "Any connection" is 0.305 as compared to 0.270 in the baseline. We have also re-estimated the model separately for new connections (appear on the second displacement) vs. disappearing connections (existing only on the first displacement) and the effects are very similar (0.26 vs. 0.32) and statistically indistinguishable suggesting that the effects of a connection to a given establishment goes to zero once

[41] For all later analyses, separate results depending on establishment size are available from the authors on request. Section 5.2 not only examines the role of group size in more detail (i.e., for the workplace, the school, and the neighborhood) but also studies other proxies of tie strength such as time since the connection was established (e.g., since they went to school together) and duration of interaction (e.g. how long they worked together) with consistent results throughout.

this connection has left.[42] Results in Section 5.2 also show that time-since-interaction reduces the estimated impact of connections. This suggests that time-invariant preferences for certain establishments are unlikely to be driving our main results.

Robustness to variations in controls and samples. Robustness checks where we vary the control set, focus on small closures, and only use cases where a single worker is connected are found in the Appendix, Table B.1. Further heterogeneity analyses in terms of person and establishment characteristics are discussed in Appendix C and D.[43] Our baseline results should be robust to demand shocks generated by the closure of an establishment since all workers (with or without a connection) at the same closing establishment should be similarly affected by such shocks. But to fully eliminate this "demand-shock" interpretation, we have verified that our results are virtually identical when we focus on connections into establishments in a different industry from that of the closing establishment.[44]

Robustness to within-worker identification. As discussed in Section 2, we are also able to estimate the relative impact of various connections using within-worker and displacement event identification where we exploit cross-sectional variation within the network of a given worker at the time of displacement (hence, here we do not require multiple displacements for the same worker). This model allows us to assess the robustness of the estimated *relative* importance of different types of connections. In the right-most column of Table 5, we present the results from estimating this model, comparing the *relative* importance of each connection, conditional on fixed-effects for the set of establishments that each displaced worker is connected to (i.e., equation 6). Said differently, these estimates are computed "within" each displaced worker's set of connections. The results mostly concur with the baseline estimates presented in the left-side column. A family member or a former co-worker clearly has a larger effect than a neighbor or a former classmate.

Replacing the connected establishment by a non-connected establishment. To verify that the estimates we find in the main analysis are unlikely to represent sorting based on market-level factors we provide two alternative test where we replace the connected establishments with other, similar, local establishments in the spirit of Kramarz and Skans (2014) and Hensvik and Skans (2016). In the first exercise, we use potential hiring establishments that are part of multi-establishment firms with at least two establishments within the same location and industry (typical cases would be, e.g., retail stores

[42] Since we require that the establishment exist during both events, establishments that did hire the worker during the first event can only be used during the second event if the worker moved before this event (as all our displacements are closures). This makes the analysis of emerging connections cleaner as the probability of entering these establishments during the first event will be close to zero.

[43] The results also remain largely unchanged if we instead estimate a model with unrestricted fixed-effects for the potentially hiring establishment (k). Note, however, that estimating this model requires that we use the same sample as in the (j,k) model, and this sample is no longer constructed from the principle of using all dyads with variations within the fixed-effects when using k-fixed-effects. The reason is that all possible combinations of workers and employers (with at least one connection) in the economy contain variation within the fixed-effects when using k-fixed-effects.

[44] See Appendix Section C for detailed results.

or restaurants) and replace the actual establishment with a randomly selected alternative within this set. Clearly, since connections may also matter at the firm, and not only the establishment, level, this exercise is somewhat stacked against our strategy if interpreted as a pure "placebo". In the second analysis, each potential hiring establishment is replaced by another establishment within a different firm, but within the same location and 3-digit industry, as a way to simultaneously examine the "common preferences" and the "demand shock" concerns. The estimates from these two "placebo-style" regressions, presented in the middle and rightmost column of Table 7, have magnitudes that are an order of magnitude smaller than the corresponding estimates of our main analysis (cf., leftmost column of Table 7). This implies that a worker who is displaced is 10 times more likely to enter an establishment where one of his or her connections are working, than into another establishment within the very same firm and location. We find this difference reassuring; if unobserved characteristics were driving our baseline results, we should instead find similar estimates for these unconnected establishments within connected firms (assuming that the production function has a large firm-level component). Similarly, if the baseline results arose because of reduced competition on the product market, rather than the connection per se, we would expect to see estimates of similar magnitudes when using unconnected establishments within the same industry, but here the differences are even smaller than in the within-firm exercises.

Table 6: Estimates for other non-connected potential hiring establishments

	Baseline[a]		...same firm[b]		...same industry[c]	
	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)
Panel A:						
Any connection	0.270	(0.005)	0.029	(0.004)	0.014	(0.001)
Constant	0.026	(0.000)	0.006	(0.001)	0.006	(0.000)
Panel B:						
Family member	1.095	(0.020)	0.037	(0.011)	0.016	(0.004)
Former co-worker	0.253	(0.010)	0.075	(0.012)	0.016	(0.003)
Former classmate	0.066	(0.004)	0.011	(0.004)	0.010	(0.002)
Current neighbor	0.086	(0.008)	0.015	(0.009)	0.023	(0.006)
Constant	0.027	(0.000)	0.006	(0.001)	0.006	(0.000)
No of fixed-effects		2,087,560		391,438		1,540,941
No of observations		41,111,774		3,676,175		29,891,982

Notes: Data are in dyadic form with one observation per combination of displaced worker and potential hiring establishment. All estimates are expressed in percentage points, i.e., the coefficients are multiplied by 100. All estimations include establishment-pair(-and-year) fixed-effects. Standard errors are clustered on the potential hiring establishment-and-year level.
[a] Repeats the first column of Table 5.
[b] Each potential hiring establishment has been replaced by another randomly selected establishment within the same firm, location (i.e., municipality), and industry (i.e., 3-digit code).
[c] Each potential hiring establishment has been replaced by another randomly selected establishment within the same location (i.e., municipality) and industry (i.e., 3-digit code).

4.3 The causal impact, by person and establishment effects

In order to estimate if and how the causal effects of connections affect sorting inequality, we examine the interaction between having a social connection and the estimated (AKM) person and establishment effects. The analysis thus relies on the model outlined in equation (4). All empirical models control for the (AKM) types of the agents and the types of social connections within their social networks. As noted in Section 2, the person effects are estimated using only the observations that precede job displacement to avoid reverse causality (e.g., being hired through connections may lead to higher wages),[45] but also as a response to the so-called limited mobility bias (see our preceding discussion).[46]

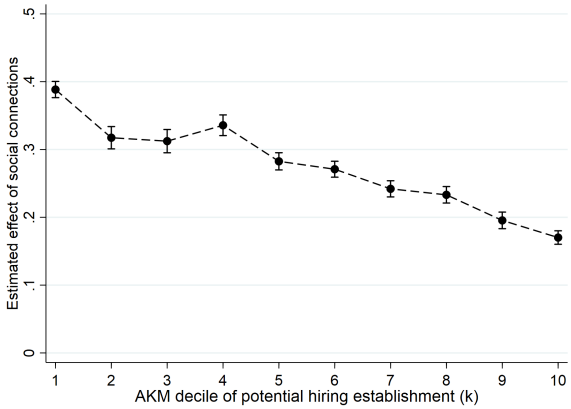


Figure 2: Estimated effects of any social connection by deciles potential hiring establishments' AKM effects (i.e. ψ_k)

Notes: The figure shows interactions with AKM deciles of potential hiring establishment (k) using the baseline model with j,k-fixed-effects. All estimates are expressed in percentage points.

We first focus on the role of establishment and person effects separately before turning to their interplay. Estimates of the effects of social connections across the distribution of demand-side heterogeneity shown in Figure 2 show that this dimension is highly relevant. The impact of social connections is more than twice as large for low-wage establishments than for high-wage establishments. The relationship in-between is approximately linear.[47]

[45] For evidence in this direction, see Appendix Table 9.

[46] Below, we verify that the results are robust if we use alternative ways of classifying firm groups.

[47] Estimates in Figure 2 and corresponding estimates for persons in Figure 5 below are from separate regressions but results are identical when estimated jointly.

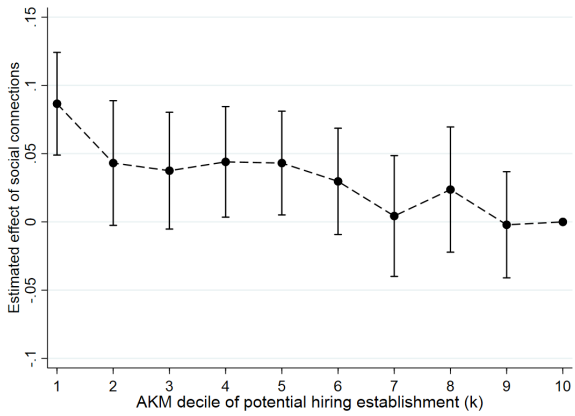


Figure 3: Within-worker identification estimates by AKM decile of potential hiring establishment (ψ_k)

Notes: The figure is based on the model of equation (6), which includes worker and set of connected establishments and year fixed-effects. It shows the interactions with AKM deciles of potential hiring establishment. Decile 10 is the reference category. All estimates are expressed in percentage points.

However, if workers who are more likely to use connections also are more likely to be connected to low-wage employers, our estimates would reflect this unobserved heterogeneity. Fortunately, using within-worker variation allows us to assess this concern. We thus estimate the individual (times connected set) fixed-effects model of equation (6) to see if workers with connections to both high- and low-wage establishments are more likely to "use" their connections to low-wage establishments. The reference point here is a connection to an establishment in the highest decile of AKM-effects. Indeed, as shown in Figure 3, connections to low-wage employers have larger causal effects, even within a given worker. Therefore, the resulting structure of the estimates from the within-worker model is very similar to that of the baseline j,k-fixed-effects model, despite the differences in identifying variation: connections to low-wage establishments have larger causal effects than connections to high-wage establishments.

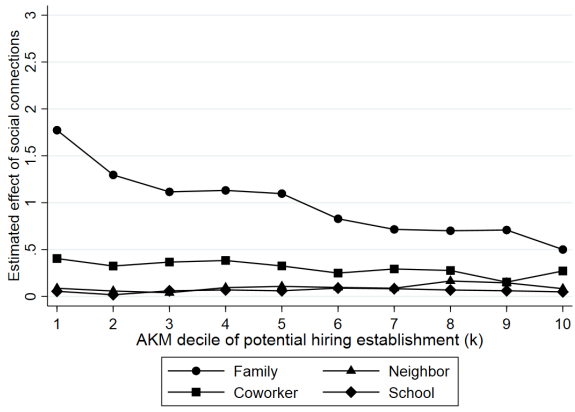


Figure 4: Estimated effect of each type of social connection by decile of the displaced potential hiring establishments' AKM effects (i.e. deciles of ψ_k).

Note: The figure shows interactions with AKM deciles of potential hiring establishment (k) using the baseline model with j,k-fixed-effects. For coefficients and standard errors of corresponding estimated linear slopes, see Table B.3 in the Appendix. All estimates are expressed in percentage points.

Estimates of the impact of connections across the distribution of demand-side heterogeneity (AKM decile of potential hiring establishment) separately by type of connection (i.e., family member, former co-worker, former classmate, or current neighbor), are graphically presented in Figure 4. The graph clearly shows that the effects of family members are more than twice as large when connecting to low-wage establishments as when connecting to high-wage establishments. This pattern is closely mirrored also for coworkers; the negative slope for these connections is less visible in the graph because the mean effect is much lower, but using a model with a linear slope (see Appendix Table B.3) provides a statistically significant estimate of -0.018 which implies that the effect in low-wage establishments is almost twice as large as in high-wage establishments also for co-worker connections.[48] The effects of former classmates and current neighbors are so limited in magnitude that we cannot identify any significant heterogeneity across establishment types.[49]

[48]Recall that the mean effect of co-workers is 0.253, and the difference (as implied from the linear model) between lowest and highest decile is 0.18.

[49]See Appendix Table B.3 for details.

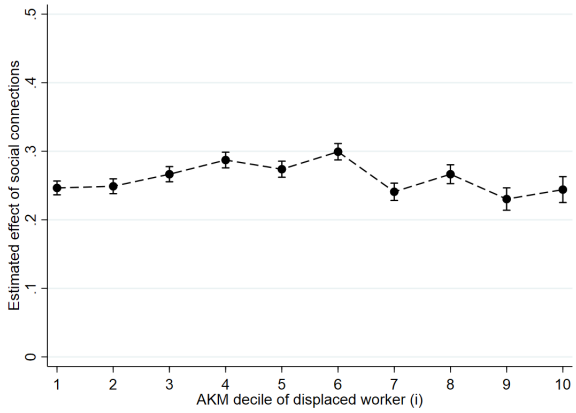


Figure 5: Estimated effects of any social connection by deciles of the displaced workers' AKM effects (i.e. θ_i)

Notes: The figure shows interactions with AKM deciles of displaced worker (i) using the baseline model with j,k-fixed-effects. All estimates are expressed in percentage points.

Estimates of the effects of social connections across the distribution of displaced person-effects (i.e., supply-side heterogeneity) are presented in Figure 5. The estimates show, in sharp contrast to the demand side heterogeneity, that connections appear to have equal causal effects for low- and high-wage displaced workers.[50] If anything, the impact appears slightly larger in the middle of the person-effects distribution.

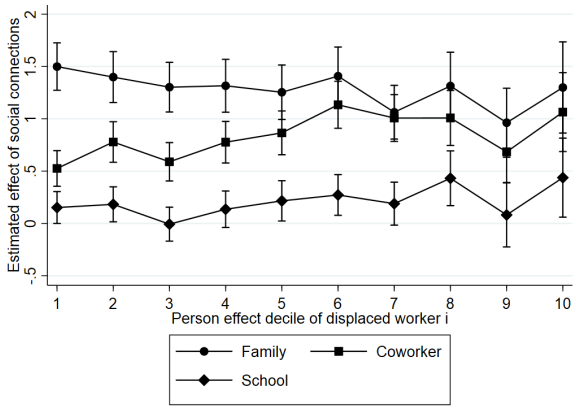


Figure 6: Individual fixed-effects estimates of the effect of social connections relative to current neighbors by AKM effect (θ_i) decile of displaced worker

Notes: The figure is based on the model in column (3) of Table 4, which includes individual and set of connected establishments and year fixed-effects. It shows the interactions with connection type and the AKM person effect decile. Current neighbors is the reference category. Standard errors are clustered on the potential hiring establishment-and-year level. All estimates are expressed in percentage points.

Within-worker identification allows us to assess whether the relative importance of the types of connections for a given worker is stable across the distribution of person effects.[51] In Figure 6 we show the

[50] In Appendix Figures B.4 and B.3 we show that the relationship is the same if we use person-effects after conditioning on age, gender, and education level.

[51] In the Appendix (Table B.3) we show linear relationships with their standard errors from the baseline model with j,k fixed-effects using both raw and residualized person effects.

relative impact of family members, former classmates, former co-workers, and current neighbors across the distribution of person effects. The Figure uses current neighbors as the baseline category across the distribution. The point estimates show that the ranking of effects (i.e., family first then co-workers, classmates, and neighbors [normalized to zero] last) is the same across the distribution of person effects. The magnitudes of the differences are, however, largest (and statistically significant) for low-wage workers and less pronounced (and not significant) at the very high end. Low-wage workers are 1.5 percentage points more likely to be hired by establishments where they have family members than where they have neighbors, whereas the corresponding difference for co-workers is 0.5 percentage points. The effects are less dispersed for high-wage workers, because family members are less important for these workers and co-workers instead matter more. The impact of former classmates are only marginally more positive than current neighbors across the distribution of person effects.[52]

Finally, we turn to the interaction between supply- and demand-side heterogeneity. Here, we use a second-order polynomial of both (AKM) worker and establishment effects as outlined in equation (5). We illustrate the estimates graphically in Figures 7a to 7d, which all present different aspects of the same regression. We report the parameter estimates and standard errors in the Appendix Table B.4 for completeness.

The effects of a connection vary with the type of worker, separately for different types of connected establishments, are displayed in Figure 7a. The effects of connections are larger for low-wage establishments regardless of whether the worker is a high-wage (dotted lined) or a low-wage (solid line) individual. Slopes are only marginally different across types of workers and the interaction is not statistically significant.See Appendix Table B.4).

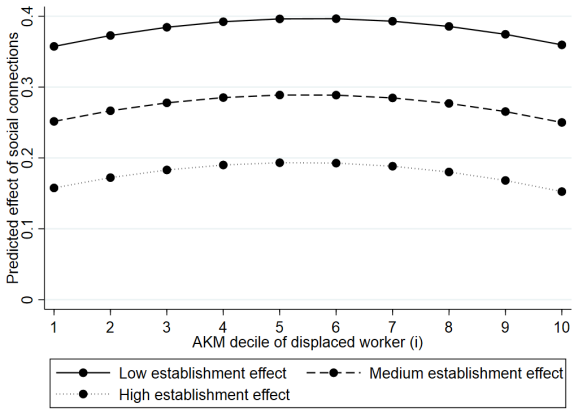
Next, we show how the causal effects of a connection vary with the type of worker, separately for different types of connected establishments (using the same underlying estimates). As Figure 7b demonstrates, the effect of connections to establishments with the lowest AKM-effects is twice as large as the effect of connections to establishments with the highest AKM-effects *regardless* of whether the displaced worker is a low-, medium-, or high-wage worker.

It is well-known that the estimated covariance between person and firm effects, in general, is negatively biased as described in detail in, e.g., Andrews et al. (2008) and Kline et al. (2018). The first-order impact of this bias, arising because any over-estimation of establishment effects immediately is transposed into a negative bias in person-effects of that establishment's employees, is removed in this setting because we correlate the person effects (estimated with pre-displacement data only) with the establishment effects of *future* potential employers. But to ensure that our estimates are not affect by the overall

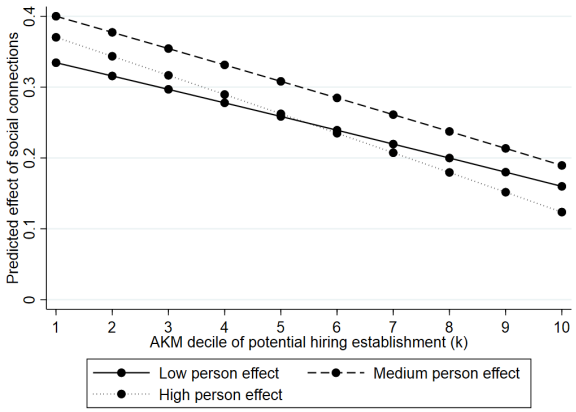
[52]Since the baseline results suggested a larger impact of former classmates from university in the individual fixed-effects model, we explored models separating between classmates from high school and university but the relationships to the person effects were in fact very similar.

noisiness of AKM-estimates for small employers, we have verified that these results are valid also if using other strategies for classifying employers. We first classified all establishments using the clustering-approach of Bonhomme et al. (2018) where we cluster the establishments (treating each year separately) into 10 groups based on the earnings distribution and then rank the clusters based on mean earnings as in Bonhomme et al. (2018). Person effects are estimated as fixed effects, using pre-displacement data only, in wage regressions that control for cluster-group fixed effects and the same observable covariates as in the AKM-model. We also grouped establishments into 10 groups based on firm-level value added per worker and proceeded as above. We then re-estimated the interacted model and derived figures corresponding to Figure 7a and Figure 7b. The ensuing results are presented in appendix Figure 8. The results show that connections to high-wage (high productivity) employers matter the most for all types of workers regardless of the method we use.

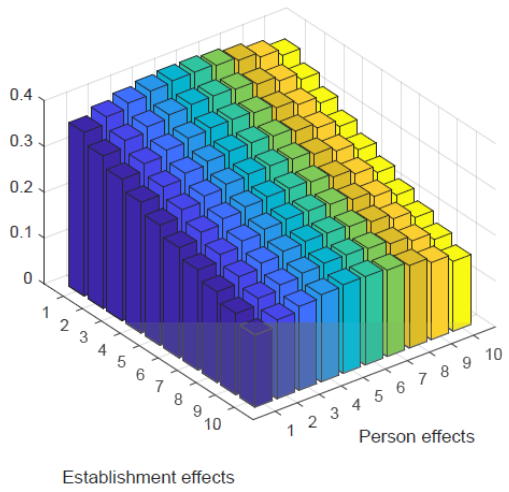
(a) By AKM decile of displaced worker (i) for low, medium and high terciles of establishment (k) effects



(b) By AKM decile of potential hiring establishment (k) for low, medium and high terciles of displaced worker (i) person effects



(c) By joint distribution of person (i) and establishment (k) AKM deciles



(d) Baseline hiring probabilities without connections, by joint distribution of person (i) and establishment (k) AKM deciles

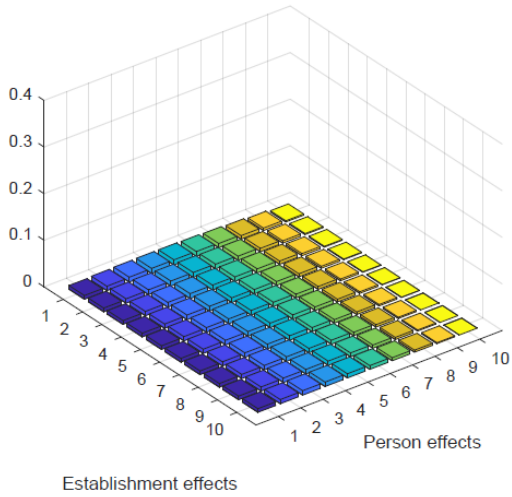
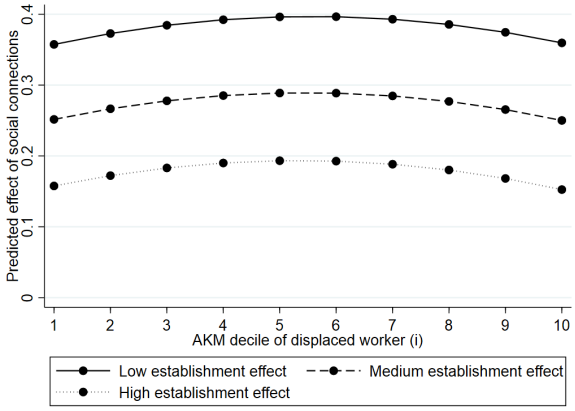


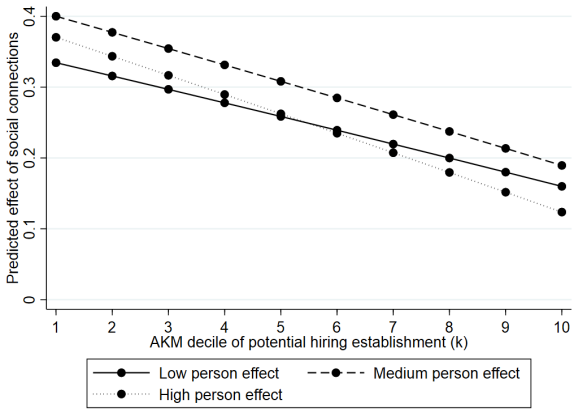
Figure 7: Effects of social connections on hiring, by joint distribution of person and establishment effects

Notes: The figures show the predicted hiring effect of social connections obtained from estimating equation (4), i.e., from interacting social connections with a second order polynomial of both the AKM person effect of the displaced and the establishment effect of the potential hiring establishment (in percentiles). All estimates are from the same regression. Panel a) and b) show slices of the 3d graph of panel c). For underlying estimates and standard errors see the Appendix Table B.4. All estimates are expressed in percentage points.

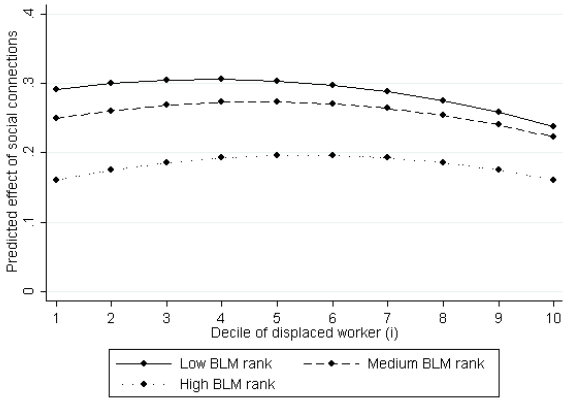
(a) By AKM decile of displaced worker (i) for low, medium and high terciles of establishment (k) effects (as in Figure 7a)



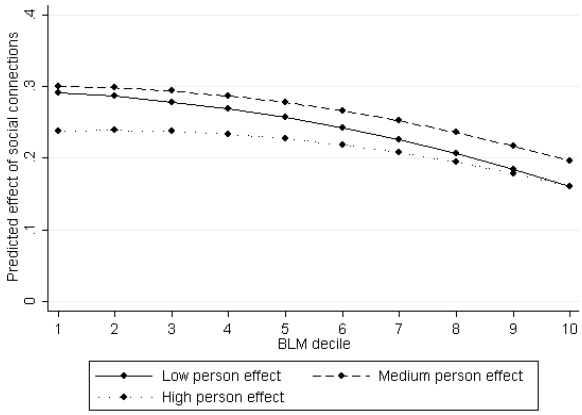
(b) By AKM decile of potential hiring establishment (k) for low, medium and high terciles of displaced worker (i) person effects (as in Figure 7b)



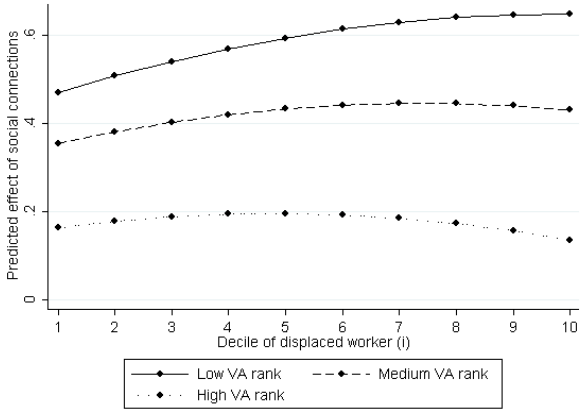
(c) Establishment ranks from BLM clustering



(d) Establishment ranks from BLM clustering



(e) Establishment ranks from VA-distribution



(f) Establishment ranks from VA-distribution

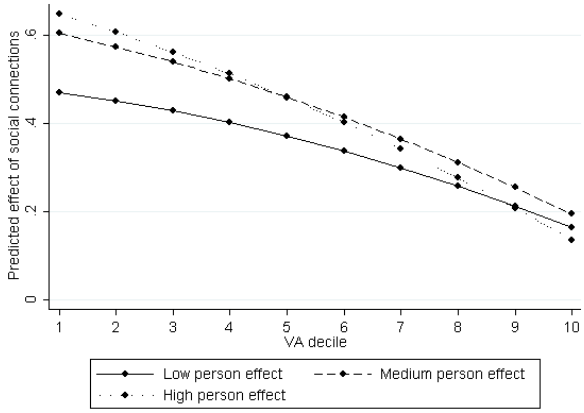


Figure 8: Alternative measures of establishment deciles

Notes: The figures show the predicted hiring effect of social connections obtained from estimating equation (4), i.e., from interacting social connections with a second order polynomial of both the person effect of the displaced and the establishment effect of the potential hiring establishment (in percentiles). Panel a) and b) is the same as panel a) and b) of Figure 7. In panel c) and d) use BLM clustering to rank establishments, while the worker deciles are derived from a wage regression similar to eq. (3), but replacing the establishment effects with BLM ranks. In panel e) and f) we repeat this exercise when ranking establishments based on value added. All estimates are expressed in percentage points.

Figure 7c shows the full 3D-graph of causal effects as functions of the joint distribution of person and establishment effects (as approximated by the second-order polynomial). Again, the Figure shows that all effects are much larger for low-wage establishments, *regardless* of the person effect. Hence, the distribution of causal effects of connections (i.e., the $\gamma(\theta, \psi)$ in equation (4)) does not contribute to sorting inequality. For completeness, we also show in Figure 7d how the person and establishment effects interact with the probability to be hired for non-connected workers; these effects are small throughout.

To ensure that the results are not reflecting underlying individual heterogeneity, we use within-worker identification, which is immune to such concerns, to analyze the relationship between the causal effects of connections and the person and establishment effects of the agents. When we use this model, the presence of individual fixed-effects limits identification to parts of the interacted second-order polynomial (see equation (7)). Results are again illustrated graphically in Figure 9 with point estimates and standard errors deferred to Appendix Table B.4. The results show that a given (displaced) worker is more likely to be hired thanks to connections to a low-wage establishment, regardless of whether she is a low-, medium-, or high-wage worker. The triple interaction between having a connection and the (AKM) person and establishment effects is positive, but tiny in magnitude (as is obvious from the graph) and not statistically significant.[53] Thus, the results using within-worker identification are in full agreement with those of the baseline model; causal effects are always larger for low-wage establishments, independently of the person-effect.

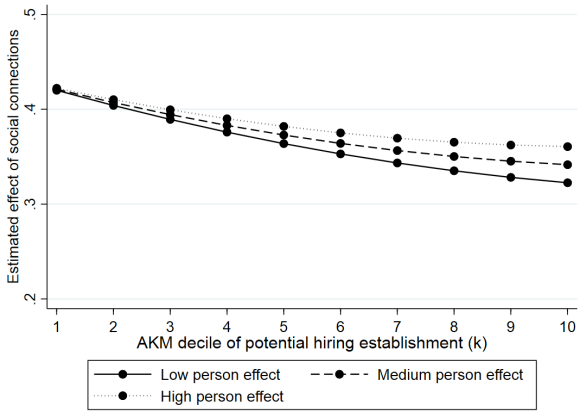


Figure 9: Within-worker identification estimates by potential hiring establishment deciles, for low-, medium-, and high-wage displaced workers.

Notes: The figure is based on the model of equation (6), which includes worker and set of connected establishments and year fixed-effects. It shows the interactions with a second order polynomial of the AKM effect of the potential hiring establishment and the interaction between the AKM person effect of the displaced and the establishment effect of the potential hiring establishment. Note that the baseline person effect of the displaced is absorbed by the (worker) fixed-effect. The interaction term between person and establishment effect is not significant. (see Appendix Table B.4). All estimates are expressed in percentage points.

[53] See Appendix Table B.4.

4.4 Social connections and sorting patterns

As shown in the previous sub-sections, social networks exhibit homophily: high-wage workers are connected to other high-wage workers. But our analysis also demonstrates that the hiring power of a social connection is as large when it connects a high-wage worker to a low-wage establishment as when it connects a *low*-wage worker to a low-wage establishment. Thus, the structure of social networks adds to sorting inequality, whereas the interactions between AKM-components and the magnitudes of the causal effects appear to by unrelated to sorting inequality. Below we assess if the combination of these two forces make connected hires more (or less) sorted than market hires.

To answer this question, we investigate how the person effects of newly hired workers co-vary with the AKM effect of the hiring employer. This analysis presents estimates that require more of a leap-of-faith to be interpreted as causal since the sample is selected to be those that actually find employment and since the identification strategy lacks the granular fixed-effects controls of the dyadic model. In addition, this strategy will more clearly suffer from attenuation bias because we cannot measure all possible friendship relations as discussed in Section 3.3.1 and since our non-connected, "market", matches therefore almost certainly include cases where unobserved social relations played a role in the process. On the other hand, this analysis also lends itself to an extension to all movers (i.e., also those that are not displaced). These non-displaced movers are analysed to ensure that the patterns we describe are not limited to the sample of displaced workers.

We estimate the following model on the sample of newly hired workers (an observation is a hired worker):

$$\theta_{ik} = \mu + X_i\beta + gC_{ik} + s\tilde{\psi}_k C_{ik} + m\tilde{\psi}_k(1 - C_{ik}) + e_{ik}, \tag{8}$$

where $\tilde{\psi}_k$ indicates (for ease of exposition) deviations from within-sample means of ψ_k. As in the rest of our analysis, θ_{ik} is constructed from an AKM-model using only pre-hire data. The parameter g captures the mean "effect" of being hired through a social connection, s captures differences in hiring patterns of socially connected workers between high- and low-wage establishments, and m captures similar differences in hiring patterns for non-connected workers ("market" matches). Differences between estimated s and m indicate differences in sorting between connected and non-connected hires. To reduce the impact of potential differences in observable characteristics, X include the establishment effect of the previous employer ($\tilde{\psi}_j$) and observable worker characteristics (see Table notes for details). As mentioned above, the model can only be estimated for the sample of realized hires.

Table 7: AKM person effect of new hire as a function of the AKM establishment effect of the hiring establishment and social connections

	(1)		(2)		(3)		(4)	
	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)
Panel A: Hired displaced workers								
Connected hire	1.467	(0.500)	1.573	(0.499)	1.612	(0.499)		
Hiring establishment AKM-effect								
× Connected hire	0.048	(0.008)	0.032	(0.008)	-0.027	(0.009)		
× Market hire	0.079	(0.003)	0.064	(0.003)	0.008	(0.005)		
F-test: Interactions are equal (p-val.)		0.0002		0.0001		0.0000		
No of observations		83,540		83,540		83,540		
R-squared		0.264		0.267		0.269		
Panel B: All hires								
Connected hire	2.301	(0.084)	2.524	(0.084)	2.481	(0.084)	1.479	(0.094)
Hiring establishment AKM-effect								
× Connected hire	0.081	(0.001)	0.051	(0.001)	-0.009	(0.001)	-0.020	(0.002)
× Market hire	0.116	(0.000)	0.088	(0.001)	0.029	(0.001)	–	–
F-test: Interactions are equal (p-val.)		0.0000		0.0000		0.0000		–
Observations		3,315,521		3,315,521		3,315,521		3,315,521
R-squared		0.302		0.306		0.309		0.408
Year specific effects		Yes		Yes		Yes		Yes
Origin establishment AKM-effects		No		Yes		Yes		Yes
Hiring establishment fixed effects		No		No		No		Yes

Notes: We account for the age, gender, education level and the number of connections of the job mover.

Table 7 displays the estimation results. The estimates show that the average "impact" of social connections is positive, which implies that employers, on average, hire more high-wage workers through social connections (as postulated by Montgomery (1991)). Furthermore, we see that job-to-job transitions are associated with sorting inequality, i.e., high-wage employers are more likely to hire high-wage workers regardless of whether they hire through social connections or through the "market". Thus, hired workers are positively sorted in terms of wage potential. However, this sorting is (statistically) significantly *less* pronounced among workers hired through connections than among workers hired through the market. These patterns arise not only in the sample of hired displaced workers (Panel A) but also in the full sample comprising all job-to-job movers (Panel B).

To illustrate the role of connections for the different types of establishments, Figure 9 uses the estimates of Table 7 to trace out the person effects as a function of the AKM establishment effects for connected hires and market hires. Figure 9a, "All hires", clearly shows that low-wage employers hire better workers through connections than through the market, whereas the converse is true for high-wage employers, Figure 9b shows a similar pattern for displaced hires. Interestingly, comparing Figures 9a and 9b, sorting is steeper for the former (All hires) than for the latter (Hired displaced workers) suggesting that our identification strategy helps us attenuate the component of sorting due to the option to stay in one's origin establishment. In the Appendix Figure B.9, we show results from a corresponding

analysis where we use person effects estimated from the full sample (instead of the pre-displacement sample) with very similar results.

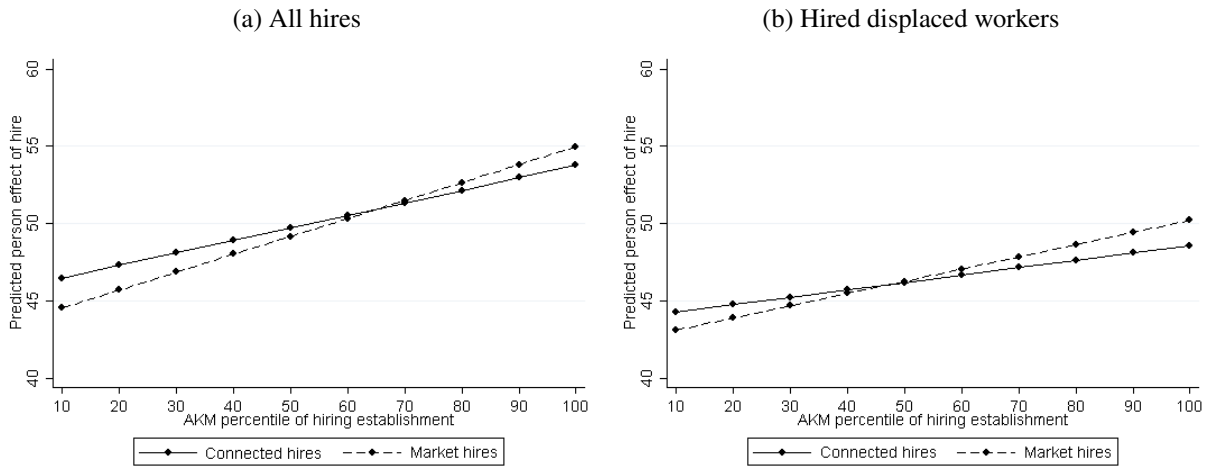


Figure 10: Predicted person effect of new hire as a function of hiring establishment effects and the use of social connections

Notes: The figure shows the predicted relationships between person effects of new hires, the hiring establishment effect and the hiring channel (connection/no connection). For estimates and standard errors see Table 7. We have added the mean person effect within each sample.

5 Validation Analyses

In the coming subsections, we show how connections operate in a number of dimensions related to our main analysis. The exercises lend further support for our claim that the main results are due to social interactions between measured connections, thus validating our approach and interpretation. These subsections are also intended to demonstrate the robustness of our previous findings in a broad set of dimensions.

5.1 Competing connections and their quality

The results from the previous Section demonstrate that connections are useful for establishments in their hiring process. If connections indeed matter, then a "better" connection should be more useful than a worse one. In addition, if connections matter and multiple connections coexist, they may affect each other.

Indeed, establishments may face a "choice" between multiple connected workers and therefore have to select the connected worker that fits them best. This idea, that connections are competing, has a long tradition in the literature on job search networks (see, e.g., Boorman, 1975; Calvó-Armengol and Jackson, 2004). To study how competition operates, we analyze what happens when multiple displaced workers are connected to the same establishment. We test two aspects of these competition effects: (i)

when competing connections co-exist, and (ii) when the relative quality of these competing connections differ.

If an establishment is connected to several displaced workers through its employees, competition should reduce the probability that a given connection results in a hire. To test this hypothesis we estimate the following extension of our main model:

$$H_{ijk} = \alpha_{jk} + X_{ik}\beta + \gamma C_{ijk} + \lambda C_{ijk} B_k + \varepsilon_{ijk}, \tag{9}$$

where B_k is an indicator for multiple connections of displaced workers to k in the same year.[54] The results are presented in Panel A of Table 8. The resulting estimate indeed shows that the causal effect of a connection is much lower if other displaced workers are connected to the same establishment. Results are similar if we instead use the number of competitors or the number of competitors as a share of all incumbent employees as our competition measure. We show separate estimates depending on whether the connected establishment (k) is low, middle, or high ψ_k. The usefulness of connections is largest for low-ψ_k establishments, whereas the negative role of competitors is growing in ψ_k, thus competition is more important at more attractive employers.

If an establishment is connected to several displaced workers through different employees, the relative quality of these competing connections should affect who will be hired. Previous studies that have investigated the role of network quality for the re-employment of displaced workers have measured quality in terms of the employment rate in the particular network. In contrast to this literature, we try to predict the precise destination of displaced workers using the a direct measure of the quality of the network as provided by the estimated person effects of these intermediary workers (θ_l). Hence, Panel B, present results from an extended version of equation 9 that include an indicator for cases where the displaced worker faces competition from another displaced worker endowed with a "better" (in terms of person effects) connecting intermediary worker. The results, consistent across the high-wage/low-wage status of the establishment, imply that competition of better-connected workers reduces the predictive hiring power of a worker's social connection.[55] We have also explored the role of the relative quality of the displaced workers (in terms of person effects), but the results show that this aspect plays no consistent role.[56]

[54] The year aspect does not matter in other parts of our analysis as displacements jk by definition are year-specific.

[55] We find similar results if we instead measure competing quality based on the observable (combined) characteristics of the displaced worker, the type of connection and the intermediary worker.

[56] Thus, connected workers are not crowded out (more) if competing connected workers are of "better" quality. This result is reasonable in light of our main analysis which suggested that "better" and "worse" workers are equally likely to be hired by the connected establishment, see Figure 3.

Table 8: The role of competing connections (of higher quality), by potential hiring establishment effects

	Estimated AKM effects of (potential) hiring establishment					
	Low ψ_k		Medium ψ_k		High ψ_k	
	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)
Panel A						
Any connection (C)	0.2513	(0.0093)	0.2127	(0.0069)	0.1628	(0.0059)
C × Competing Connection (CC)	-0.0691	(0.0185)	-0.1022	(0.0182)	-0.1301	(0.0191)
Constant	0.0164	(0.0005)	0.0210	(0.0004)	0.0149	(0.0003)
No of fixed-effects		478,128		684,111		734,016
No of observations		9,496,877		13,182,356		13,747,457
Panel B						
Any connection (C)	0.2604	(0.0098)	0.2112	(0.0069)	0.1562	(0.0059)
C × CC	-0.0196	(0.0220)	-0.0568	(0.0213)	-0.0573	(0.0229)
C × Higer CC Quality	-0.1038	(0.0225)	-0.0699	(0.0162)	-0.0964	(0.0151)
Constant	0.0163	(0.0005)	0.0210	(0.0004)	0.0149	(0.0003)
No of fixed-effects		478,128		684,111		734,016
No of observations		9,496,877		13,182,356		13,747,457

Notes: The regressions are restricted to potential hiring establishments where no two (or more) employees had connections to the same displaced worker. Observations with missing quality measures are accounted for missing variable indicators. All estimates are expressed in percentage points, i.e., the coefficients are multiplied by 100. All estimations include establishment-pair(-and-year) fixed-effects. Competition and quality measures are demeaned (using the mean among those with a connection). Standard errors are clustered on the potential hiring establishment-and-year level. "Competing Connection" is an indicator taking the value one if some other displaced worker (in the same year) has a connection to the same establishment. "Higher Competing Connection Quality" (HCCQ) is an indicator variable taking the value one if (some) Competing Connection is mediated by an intermediary worker with a higher estimated person effect than subject i's intermediary worker (i.e., $HCCQ_{ik} = I[\theta_{l(\tilde{i},k)} > \theta_{l(k,i)}$ for some $\tilde{i}$ who is displaced in the same year as i with a connection to k.).

5.2 Connections' strength and intensity

In this subsection, we examine how strength and intensity of social connections are related to the causal effects. The general presumption is that, if indeed social interactions are important, we should see larger effects in those cases where we expect interactions to be more frequent, i.e. in cases where the social ties are stronger.

Our main results show that a family member is associated with a much larger causal effect than any other connection in our data, even though the networks provided by family members exhibit less homophily in the AKM-dimension than professional connections. We interpret this as evidence in favor of strong social connections being particularly important on the labor market (see, e.g., Boorman (1975)).

To further test the robustness of this conclusion, we have re-estimated our model, but now allowing for further heterogeneity in a number of dimensions: the size of the group where the interaction took place (more detailed than in Table 5), the duration of interaction, and the time since interacting. These dimensions can be viewed as proxies for "ties' strength", and in particular the intensity of the interaction. Assessing the importance of these factors will also shed some additional light on the data restrictions discussed in Section 3.3 where we imposed a set of boundaries on the observed connections. For example, former high-school classmates are (for data availability reasons) not observed for

those who graduated before 1985. Moreover, as is apparent from Section 3.3, our measures of social connections are comprehensive and largely error-free in a statistical sense (co-workers were paid by the same employer), but they do not necessarily capture agents that frequently interact on a social level. Hence, we decided to focus on those cases where the connections should be well-measured and most meaningful. Most notably, we excluded networks with 100 or more individuals.

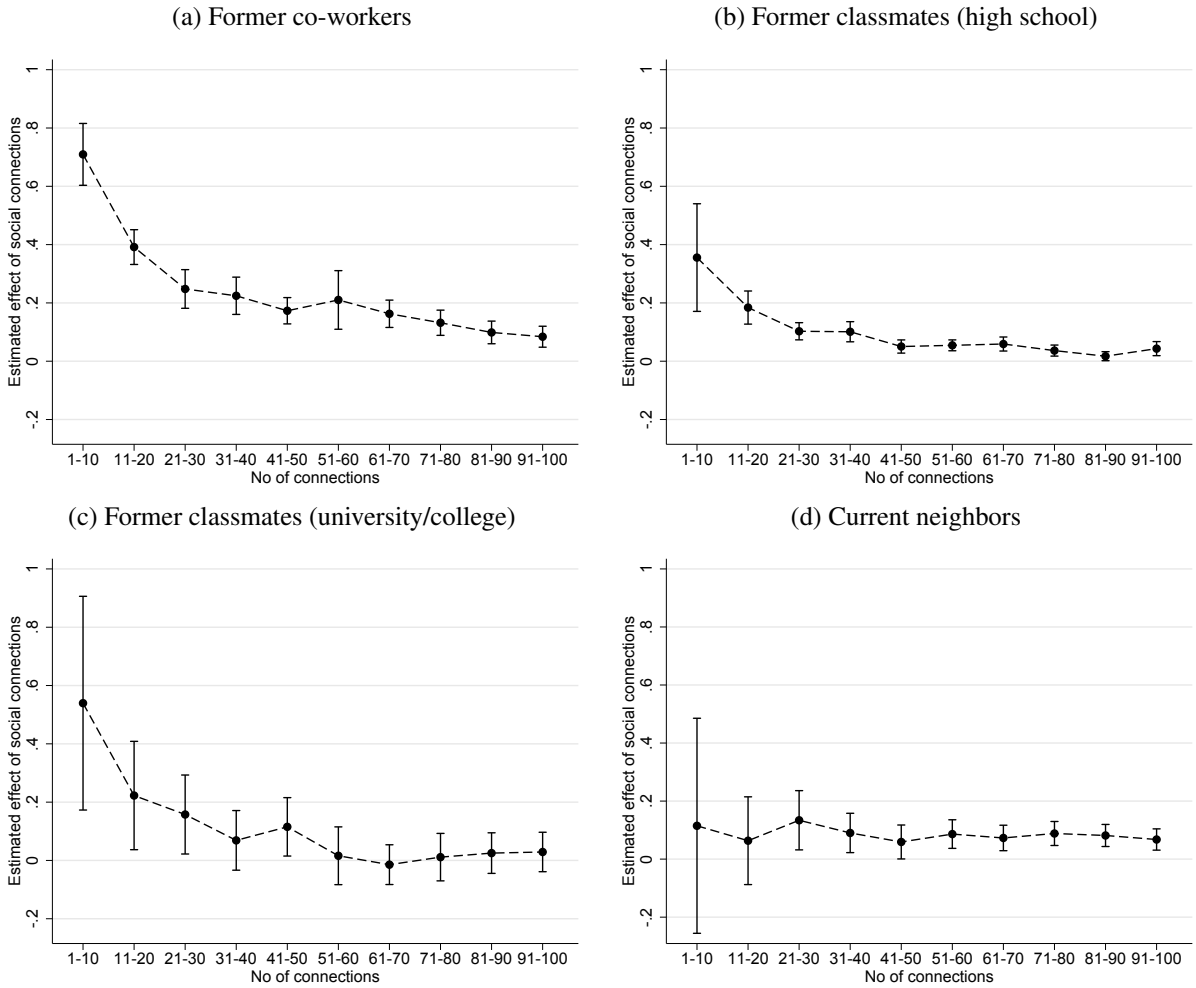


Figure 11: The estimated impact of a connection on the hiring probability, by the size of the particular network (i.e., the workplace, class, or neighborhood), with 95 percent confidence intervals, for former co-workers, former classmates (high school and college/university), and current neighbors

Notes: All estimates are obtained from the same estimation, where the indicator for the particular connection has been replaced by its interactions with the size of the particular network (10 categories). The estimates are expressed in percentage points, i.e., the coefficients are multiplied by 100. All estimations include establishment-pair(-and-year) fixed-effects. Standard errors are clustered on the potential hiring establishment-and-year level.

First, the size of the group in which the interaction took place matters. We analyse group size for three types of measured connections: former co-workers, former classmates, and current neighbors. As can be seen in Figure 11, the magnitudes of the estimates are decreasing with the size of the group; especially so among former co-workers and classmates (from both high-school and college/university). Importantly, group size matters also for groups that are already very small, co-workers from sites with 10 or less employees matter more than those with 11–20 employees, even though it is highly likely that all co-workers at an establishment with 11–20 employees know each other. This strongly suggests that connections matter and are more useful if they are formed in smaller social groups. It also supports our choice to exclude groups of more than 100 individuals since they appear to be uninformative regarding the *relevant* connections.

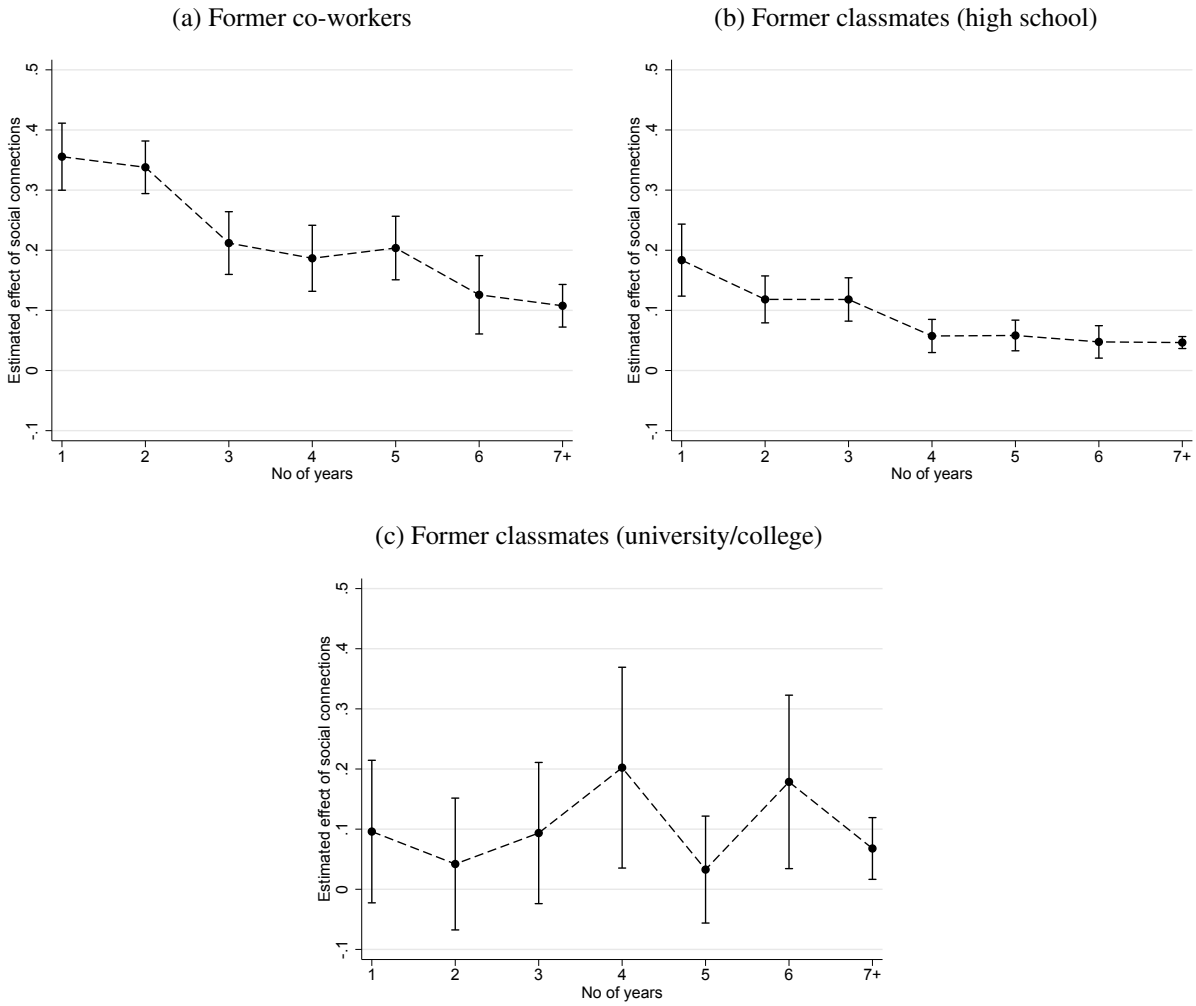


Figure 12: The estimated impact of a connection on the hiring probability, by time since interaction, with 95 percent confidence intervals, for former co-workers and former classmates (high school and college/university)

Notes: All estimates are obtained from the same estimation, where the indicator for the particular connection has been replaced by its interactions with time since interaction (7 categories). The estimates are expressed in percentage points, i.e., the coefficients are multiplied by 100. All estimations include establishment-pair(-and-year) fixed-effects. Standard errors are clustered on the potential hiring establishment-and-year level.

Second, time-since-interaction matters. The impact of previous co-workers depreciates rapidly with time since interaction. The estimates as presented in Figure 12 suggest that the causal impact is decreasing with this time since interaction for connections with former co-workers and with former classmates, at least from high school.[57] This also suggests that our choice to consider only former co-workers from the most recent of all past workplaces (before becoming employed at establishment j) is innocuous, and that our lack of data on high-school classmates prior to 1985 is likely to be of minor importance.

Third, the duration of the interaction matters. The estimates depicted in Figure 13 show that there is a strong positive relationship between time spent together as co-workers and the magnitude of the causal estimate. There is a similar tendency among those connected through current neighbors, even though it is not as marked since all estimates in this category are small.

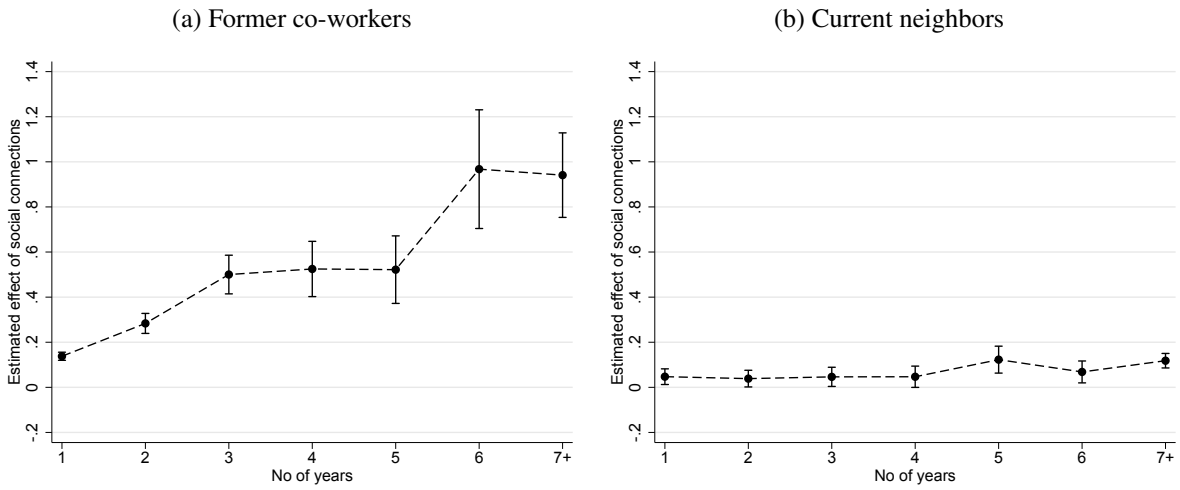


Figure 13: The estimated impact of a connection on the hiring probability, by interaction time, with 95 percent confidence intervals, for former co-workers and current neighbors

Notes: All estimates are obtained from the same estimation, where the indicator for the particular connection has been replaced by its interactions with interaction time (7 categories). The estimates are expressed in percentage points, i.e., the coefficients are multiplied by 100. All estimations include establishment-pair(-and-year) fixed-effects. Standard errors are clustered on the potential hiring establishment-and-year level.

Similarity in terms of demographic characteristics also tend to matter. We estimate if the effects are larger for agents with similar observable characteristics along the lines of Bayer et al. (2008). The results, presented in Appendix Table **??**, show that similarity in gender, age, and immigration status is more important than similarity in terms of education, providing additional support for the importance of social proximity.

Overall, these results show a consistent positive relationship between social proximity and the magnitude of the causal effects. Effects are decreasing with the size of the group where the interaction took place and with time since interaction but growing with the duration of interaction between the agents

[57] Estimates for former college/university classmates are very imprecise, and no clear result arises.

and with similarity in terms of observables.

5.3 The post-hire outcomes of connected hires

Even though the focus of the paper has been to document sorting patterns due to connections, we believe it useful to also examine if and how post-hire outcomes differ between connected hires and "market" hires. Evidence of differences is further proof that connections matter and have an impact on real outcomes.

In Table 9, we compare the earnings and employment outcomes between those with and without a connection to their new employer. The sample includes displaced workers who found a job directly after displacement. The individual-level regressions are of the form

$$Y_{ij} = \alpha_j + \lambda C_{ijk} + \beta^i X_{ij} + e_{ij} \tag{10}$$

and include closing-establishment fixed-effects, as well as controls for pre-displacement earnings and employment histories. The estimates should thus be interpreted as comparisons between similar displaced workers who found jobs with vs. without social connections.

Results show that earnings and job stability are higher for workers who were rehired through social connections. The results are near universal. The one key exception is that workers who found their next job through family members receive lower earnings early on, but positive earnings in the longer run, a result which closely mimics the results for parental contacts in Kramarz and Skans (2014). The results for earnings 3 years after for hires through university connections are very close to zero, but all other remain solidly positive in both the short and medium run. As is prevalent in the previous literature, we find positive estimates from all connections on the probability of remaining in the post-hire establishments three years after the displacement.

As a final exercise (not reported in the Table), we investigate if the relationship between the AKM-component of the hiring establishment and worker earnings differ between connected and unconnected hires. We add the hiring establishment's AKM-effect ψ_k interacted with the indicator for connections (C_{ijk}) to equation 10 and estimate the impact on log earnings of rehired worker i. The results show that workers receive similar returns from working in a better (higher ψ_k) establishment *regardless* of whether they were hired through a connection or not.[58]

[58] The estimated impact of ψ_k on log earning is 0.56 with standard error 0.03, and the interaction term is equal to 0.04 with standard error 0.09.

Table 9: Post-hire outcomes

	Outcomes after 1 year		Outcomes after 3 years					
	Log(Earnings) [a]		Log(Earnings) [a]		Employed [b]		Job stability [c]	
	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)
Panel A								
Any connection	0.083	(0.013)	0.175	(0.019)	0.029	(0.004)	0.128	(0.007)
Constant	11.043	(0.026)	10.805	(0.027)	0.747	(0.005)	0.249	(0.005)
Panel B								
Family member	0.002	(0.018)	0.152	(0.027)	0.028	(0.006)	0.136	(0.008)
Former co-worker	0.137	(0.019)	0.148	(0.031)	0.020	(0.006)	0.076	(0.011)
Former classmate	0.179	(0.029)	0.201	(0.049)	0.021	(0.010)	0.158	(0.017)
Current neighbor	0.126	(0.037)	0.197	(0.059)	0.036	(0.013)	0.152	(0.023)
Constant	11.043	(0.026)	10.805	(0.027)	0.747	(0.005)	0.249	(0.005)
No of fixed-effects		29,554		29,554		29,554		29,554
No of observations		208,738		208,738		208,738		208,738

Notes: The estimation sample is all displaced workers who were employed in in November of year $t+1$. All estimations include closing establishment (-and-year) fixed-effects and controls for the workers' age, sex, education, and three years of pre-displacement employment, earnings, and employer history. Standard errors are clustered on the closing establishment (-and-year) level.

[a] Earnings is defined as annual labor income and has been left censored at SEK 1,000.

[b] Employed is defined as being employed in November of year $t+3$.

[c] Job stability is defined as being employed at the same establishment in November of both year $t+1$ and $t+3$.

6 Conclusions

A vast number of studies have shown that social connections play a quantitatively important role in the process of matching workers to jobs. And, because social relations tend to be homophilous – with persons of similar social status being connected – the literature has presumed that social networks exacerbate labor market inequality.

In this paper, we assess this presumption. To do so, we document the causal role of a wide set of social connections (i.e., family members, former co-workers, former classmates, and current neighbors) in the matching of displaced workers and establishments within a unified empirical setting. Our analyses rely on establishment closures as exogenous events forcing workers to search for new jobs, allowing us to compare the re-employment outcomes of workers who lost their jobs in the same closure event, and to document to what extent social connections causally affect where these workers are rehired.

We use estimates from an AKM-decomposition to characterize workers and employers quality in terms of earnings capacities and discern to what extent connections affect the sorting between workers and employers. We first show that connections exhibit baseline homophily, i.e. high-wage workers are connected to other high-wage workers workers, and (to a lesser extent) also to high-wage employers. This is particularly true for professional networks such as past co-workers or classmates from university, but less so for family connections. The pattern also holds when using the residualized person effects, which only capture the "time-invariant unobservable" dimension of the person effects.

Next, we show that all our measured social connections predict post-displacement hiring patterns. By eliminating other interpretations of this finding, in particular a common preference for an establishment shared by the job searcher and the connected establishment, we believe we show that connections "cause" hiring.

The impact is largest if the social connections to an establishment is made by a family member, followed by connections through a past co-worker. Co-worker contacts are particularly useful if both worked together relatively recently (i.e., the value of the connection seems to depreciate fairly rapidly). Each former classmate and current neighbor (with children the same age) matter as well, but less than a family member or co-worker. More generally, we show a consistent positive relationship between social proximity and intensity of the connection on one part and the magnitudes of the causal estimates on the other: the more interaction time, the shorter the time since interaction, and the smaller the size of the group where the interaction took place, the more likely that the connected displaced worker is hired. These findings clearly support the presumption that measured social connections only affect hiring probabilities if social interactions are abundant.

The social connections have a larger causal impact on hiring when the demand side is a low-wage establishment, regardless of the type of worker. Displaced workers with multiple social connections are more likely to enter a low-wage establishment than a high-wage establishment if connected to both types of establishments. This is true both for low-wage and high-wage workers. As a consequence, low-wage establishments are able to attract high-wage workers through social connections.

Overall, we find that hiring through social connections is associated with *less* sorting than market matches, despite of the clear baseline homophily within our social networks and the prevalence of connections between high-wage workers and high-wage establishments. This conclusion – hiring through social connections being less sorted than market matches – holds for hired displaced workers as well as for all hires.

These novel empirical regularities stand in sharp contrast to a standard presumption shared between the theoretical literature on social networks in economics and sociology discussed in the introduction. Although it is plausible that social networks in general generate inequality of opportunities, our results clearly illustrate that this does not imply that allocations through social networks always leads to more unequal distributions than allocations through market mechanisms, even when the social network is characterized by strong homophily.

References

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

Abowd, J. M., McKinney, K. L. and Schmutte, I. M. (2018), 'Modeling endogenous mobility in earnings determination', *Journal of Business and Economic Statistics* .

Andrews, M. J., Gill, L., Schank, T. and R., U. (2008), 'High wage workers and low wage firms: Negative assortative matching or limited mobility bias?', *Journal of the Royal Statistical Society: Series A (Statistics in Society)* .

Barth, E., Bryson, A., Davis, J. and Freeman, R. (2016), 'It's where you work: Increases in the dispersion of earnings across establishments and individuals in the united states', *Journal of Labor Economics* **34**, S67–S97.

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

Bentolila, S., Michelacci, C. and Suarez, J. (2010), 'Social contacts and occupational choice', *Economica* **77**(305), 20–45.

Bonhomme, S., Lamadon, T. and Manresa, E. (2018), 'A distributional framework for matched employer-employee data', *Unpublished manuscript, University of Chicago* .

Boorman, S. A. (1975), 'A combinatiorial optimization model for transmission of job information through contact networks', *The Bell Journal of Economics* **6**(1), 216–249.

Brown, M., Setren, E. and Topa, G. (2016), 'Do informal referrals lead to better matches? evidence from a firm's employee referral system', *Journal of Labor Economics* **34**(1), 161–209.

Browning, M. and Heinesen, E. (2012), 'Effect of job loss due to plant closure on mortality and hospitalization', *Journal of health economics* **31**(4), 599–616.

Burks, S. V., Cowgill, B., Hoffman, M. and Housman, M. (2015), 'The value of hiring through employee referrals', *The Quarterly Journal of Economics* **130**(2), 805–839.

Calvó-Armengol, A., Verdier, T. and Zenou, Y. (2007), 'Strong and weak ties in employment and crime', *Journal of Public Economics* **91**(1), 203–233.

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

Calvó-Armengol, A. and Jackson, M. O. (2007), 'Networks in labor markets: Wage and employment dynamics and inequality', *Journal of Economic Theory* **132**(1), 27 – 46.

Card, D., Cardoso, A. R., and Kline, P. (2016), 'Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women.', *The Quarterly journal of economics* **131**(2), 633–686.

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

Card, D., Heining, J. and Kline, P. (2013), 'Workplace heterogeneity and the rise of west german wage inequality', *The Quarterly journal of economics* **128**(3), 967–1015.

Casella, A. and Hanaki, N. (2006), 'Why personal ties cannot be bought', *The American economic review* **96**(2), 261–264.

Charles, K. K. and Stephens, Jr, M. (2004), 'Job displacement, disability, and divorce', *Journal of Labor Economics* **22**(2), 489–522.

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

Coelli, M. B. (2011), 'Parental job loss and the education enrollment of youth', *Labour Economics* **18**(1), 25–35.

Del Bono, E., Weber, A. and Winter-Ebmer, R. (2015), 'Fertility and economic instability: the role of unemployment and job displacement', *Journal of Population Economics* **28**(2), 463–478.

Dustmann, C., Glitz, A., Schönberg, U. and Brücker, H. (2015), 'Referral-based job search networks', *The Review of Economic Studies* **83**(2), 514–546.

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

Eliason, M. (2011), 'Income after job loss: the role of the family and the welfare state', *Applied Economics* **43**(5), 603–618.

Eliason, M. (2012), 'Lost jobs, broken marriages', *Journal of Population Economics* **25**(4), 1365–1397.

Eliason, M. (2014), 'Alcohol-related morbidity and mortality following involuntary job loss: evidence from swedish register data', *Journal of studies on alcohol and drugs* **75**(1), 35–46.

Eliason, M. and Storrie, D. (2006), 'Lasting or latent scars? swedish evidence on the long-term effects of job displacement', *Journal of Labor Economics* **24**(4), 831–856.

Eliason, M. and Storrie, D. (2009*a*), 'Does job loss shorten life?', *Journal of Human Resources* **44**(2), 277–302.

Eliason, M. and Storrie, D. (2009*b*), 'Job loss is bad for your health–swedish evidence on cause-specific hospitalization following involuntary job loss', *Social science & medicine* **68**(8), 1396–1406.

Eliason, M. and Storrie, D. (2010), 'Inpatient psychiatric hospitalization following involuntary job loss', *International Journal of Mental Health* **39**(2), 32–55.

Galenianos, M. (2013), 'Learning about match quality and the use of referrals', *Review of Economic Dynamics* **16**(4), 668 – 690.

Gee, L., Jones, J. and Burke, M. (2017), 'Social networks and labor markets: How strong ties relate to job finding on facebook's social network', *Journal of Labor Economics* **35**, 485 – 518.

Glitz, A. (2017), 'Coworker networks in the labour market', *Labour Economics* **44**(Supplement C), 218 – 230.

Granovetter, M. S. (1973), 'The strength of weak ties', *American journal of sociology* **78**(6), 1360–1380.

Heath, R. (2018), 'Why do firms hire using referrals? evidence from bangladeshi garment factories', *Journal of Political Economy* **126**(4), 1691–1746.

Hensvik, L., Muller, D. and Skans, O. N. (2017), 'Connecting the young: High school graduates' matching to first jobs in booms and great recessions'.

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

Hethey-Maier, T. and Schmieder, J. F. (2013), 'Does the use of worker flows improve the analysis of establishment turnover? evidence from german administrative data', *Schmollers Jahrbuch* **133**(4), 477–510.

Hijzen, A., Upward, R. and Wright, P. W. (2010), 'The income losses of displaced workers', *Journal of Human Resources* **45**(1), 243–269.

Huttunen, K. and Kellokumpu, J. (2016), 'The effect of job displacement on couples' fertility decisions', *Journal of Labor Economics* **34**(2), 403–442.

Ioannides, Y. M. and Datcher Loury, L. (2004), 'Job information networks, neighborhood effects, and inequality', *Journal of economic literature* **42**(4), 1056–1093.

Kline, P., Saggio, R. and Solvsten, M. (2018), 'Leave-out estimation of variance components', *mimeo* .

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

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

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

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

Oreopoulos, P., Page, M. and Stevens, A. H. (2008), 'The intergenerational effects of worker displacement', *Journal of Labor Economics* **26**(3), 455–483.

Oyer, P. and Schaefer, S. (2016), 'Firm/employee matching: An industry study of us lawyers', *ILR Review* **69**(2), 378–404.

Postel-Vinay, F. and Robin, J.-M. (2002), 'Equilibrium wage dispersion with worker and employer heterogeneity', *Econometrica* .

Rege, M., Skardhamar, T., Telle, K. and Votruba, M. (2009), 'The effect of plant closure on crime'.

Rege, M., Telle, K. and Votruba, M. (2007), 'Plant closure and marital dissolution'.

Rege, M., Telle, K. and Votruba, M. (2011), 'Parental job loss and children's school performance', *The Review of Economic Studies* **78**(4), 1462–1489.

Saygin, P., Weber, A. and Weynandt, M. (2014), Coworkers, networks, and job search outcomes, Discussion Paper No. 8174, IZA, Bonn, Germany.

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

Simon, C. J. and Warner, J. T. (1992), 'Matchmaker, matchmaker: The effect of old boy networks on job match quality, earnings, and tenure', *Journal of Labor Economics* **10**(3), 306–330.

Song, J., Price, D. J., Guvenen, F., Bloom, N. and Von Wachter, T. (2019), 'Firming up inequality', *The Quarterly journal of economics* **134**(1), 1–50.

Appendix A Further descriptive statistics

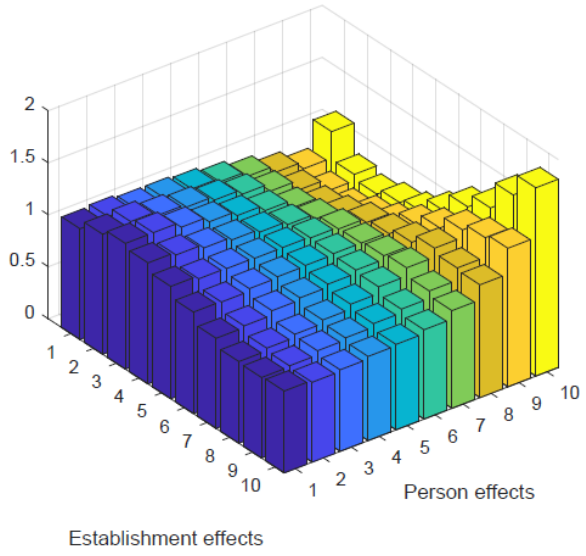


Figure A.1: Joint distribution of AKM person and establishment effects in the sample of all new hires

Table A.1: Summary statistics by closing and potential hiring establishments and for the estimation sample comprised by pairs of potential hiring establishments and displaced workers

	Closing establishments (j)		Potential hiring establishments (k)		Dyads of displaced workers (i) and potential hiring establishments (k)	
	N	%	N	%	N	%
Size						
1–9 employees	22,047	69.91	395,587	43.37	10,161,663	24.72
10+ employees	9,491	30.09	516,497	56.63	30,950,111	75.28
Age						
1–5 years	17,602	55.81	195,361	21.42	7,384,653	17.96
6+ years	13,936	44.19	716,723	78.58	33,727,121	82.04
Productivity [a]						
Low			98,408	10.79	4,750,955	11.56
Medium			151,629	16.62	7,793,214	18.96
High			153,217	16.80	9,414,202	22.90
N/A			508,83	55.79	19,153,400	46.59
AKM establishment effects						
Low	12,929	40.99	321,440	35.24	10,270,584	24.98
Medium	7,414	23.51	316,960	34.75	14,639,048	35.61
High	6,093	19.32	247,058	27.09	15,575,697	37.89
N/A	5,102	16.18	26,626	2.92	626,441	1.52
No of observations		31,538		912,084		41,111,774

[a] Productivity is only available for a subsample of all establishments.

Appendix B Further results

B.1 Further robustness

Table B.1: Sensitivity analysis; controls and samples

	Main model [a]		Excl. self-employed intermediary workers [b]		Incl. worker characteristics [c]		Incl. worker char. and AKM controls [d]		Closure size <10		At most one connection per worker and potential hiring establishment [e]	
	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)
Panel A:												
Any connection	0.270	(0.005)	0.265	(0.005)	0.270	(0.005)	0.272	(0.005)	0.360	(0.011)	0.202	(0.004)
Constant	0.026	(0.000)	0.028	(0.000)	0.028	(0.001)	0.029	(0.001)	0.065	(0.005)	0.017	(0.000)
Panel B:												
Family member	1.095	(0.020)	1.111	(0.021)	1.095	(0.020)	1.070	(0.038)	1.376	(0.049)	0.908	(0.019)
Former co-worker	0.253	(0.010)	0.256	(0.010)	0.253	(0.010)	0.252	(0.028)	0.304	(0.020)	0.149	(0.007)
Former classmate	0.066	(0.004)	0.066	(0.004)	0.066	(0.004)	0.068	(0.002)	0.064	(0.010)	0.038	(0.003)
Current neighbor	0.086	(0.008)	0.086	(0.009)	0.085	(0.008)	0.086	(0.004)	0.111	(0.028)	0.041	(0.006)
Constant	0.027	(0.000)	0.028	(0.000)	0.028	(0.001)	0.029	(0.001)	0.081	(0.005)	0.018	(0.000)
Panel C:												
Family members												
Parent	1.867	(0.052)	1.835	(0.053)	1.866	(0.052)	1.810	(0.077)	2.414	(0.125)	1.513	(0.050)
Adult child	0.670	(0.051)	0.675	(0.053)	0.672	(0.051)	0.655	(0.011)	0.760	(0.135)	0.528	(0.046)
Spouse	1.974	(0.078)	2.188	(0.086)	1.974	(0.078)	1.925	(0.011)	2.612	(0.210)	1.828	(0.080)
Sibling	0.697	(0.023)	0.702	(0.024)	0.697	(0.023)	0.693	(0.005)	0.881	(0.054)	0.524	(0.020)
Former co-worker	0.252	(0.010)	0.255	(0.010)	0.252	(0.010)	0.251	(0.003)	0.302	(0.020)	0.149	(0.007)
Former classmate												
High school	0.064	(0.004)	0.064	(0.004)	0.064	(0.004)	0.066	(0.003)	0.062	(0.010)	0.038	(0.003)
College/university	0.088	(0.018)	0.080	(0.018)	0.088	(0.018)	0.086	(0.009)	0.096	(0.065)	0.042	(0.012)
Current neighbor	0.080	(0.008)	0.081	(0.009)	0.080	(0.008)	0.081	(0.004)	0.100	(0.028)	0.040	(0.006)
Constant	0.026	(0.000)	0.027	(0.000)	0.028	(0.001)	0.029	(0.001)	0.077	(0.005)	0.018	(0.000)
No of fixed-effects	2,087,560		2,067,009		2,087,560		2,087,560		852,619		1,923,085	
No of observations	41,111,774		37,838,199		41,111,774		41,111,774		2,328,984		36,990,336	

Notes: All estimates are expressed in percentage points, i.e., the coefficients are multiplied by 100. All estimations include establishment-pair(-and-year) fixed-effects. Standard errors are clustered on the potential hiring establishment-and-year level.

[a] Repeats the estimates of the left column of Table 5.

[b] Excludes the possibility that the intermediary worker is the firm-owner.

[c] In the main estimation no (displaced) worker characteristics were included in X_{it}, while here we have included age (3 categories), sex, nativity, and attained education level (3 categories).

[d] Here we add the person effect and the interaction between the person and the establishment effect (note that the baseline establishment effect is accounted for by the fixed-effect). We dummy out the cases where there is no estimated person effect.

[e] This restriction has been imposed in the analyses in Appendix D.

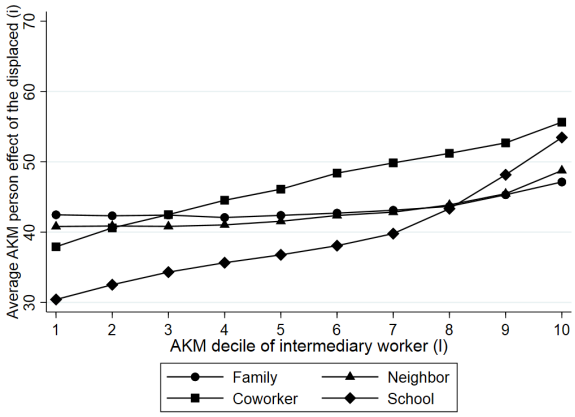
Table B.2: The estimated importance of social connections by characteristics of the potential hiring establishment (j)

	Establishment size (employees)				Industry (2 digits)				Productivity					
	1–9		10+		Same		Different		Low		Medium		High	
	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)
Panel A														
Any connection	0.368	(0.010)	0.241	(0.005)	0.585	(0.024)	0.227	(0.004)	0.320	(0.015)	0.252	(0.010)	0.212	(0.010)
Constant	0.013	(0.001)	0.031	(0.000)	0.162	(0.002)	0.015	(0.000)	0.023	(0.001)	0.020	(0.001)	0.020	(0.001)
Panel B														
Family member	1.814	(0.054)	0.896	(0.021)	3.034	(0.132)	0.948	(0.020)	1.570	(0.089)	1.182	(0.057)	0.966	(0.047)
Former co-worker	0.284	(0.018)	0.242	(0.012)	0.545	(0.040)	0.182	(0.008)	0.279	(0.029)	0.223	(0.021)	0.183	(0.020)
Former classmate	0.033	(0.005)	0.074	(0.005)	0.153	(0.023)	0.056	(0.004)	0.055	(0.013)	0.074	(0.009)	0.074	(0.009)
Current neighbor	0.060	(0.013)	0.094	(0.010)	0.358	(0.079)	0.070	(0.008)	0.117	(0.032)	0.085	(0.021)	0.118	(0.022)
Constant	0.014	(0.001)	0.031	(0.000)	0.161	(0.002)	0.015	(0.000)	0.024	(0.001)	0.020	(0.001)	0.021	(0.001)
Panel C														
Family member														
Parent	2.840	(0.127)	1.560	(0.055)	4.137	(0.313)	1.708	(0.051)	2.401	(0.203)	1.957	(0.140)	1.618	(0.121)
Adult child	1.594	(0.168)	0.452	(0.050)	2.148	(0.366)	0.560	(0.048)	1.119	(0.226)	0.583	(0.134)	0.544	(0.122)
Spouse	4.558	(0.252)	1.347	(0.075)	6.167	(0.521)	1.640	(0.073)	2.813	(0.338)	2.525	(0.243)	1.344	(0.163)
Sibling	0.972	(0.056)	0.618	(0.025)	2.146	(0.154)	0.583	(0.022)	1.077	(0.103)	0.771	(0.066)	0.717	(0.058)
Former co-worker	0.280	(0.018)	0.242	(0.012)	0.544	(0.040)	0.181	(0.008)	0.278	(0.029)	0.223	(0.021)	0.183	(0.020)
Former classmate														
High school	0.035	(0.006)	0.072	(0.005)	0.154	(0.024)	0.054	(0.004)	0.056	(0.013)	0.073	(0.009)	0.070	(0.010)
College/university	0.006	(0.012)	0.099	(0.021)	0.146	(0.066)	0.078	(0.018)	0.090	(0.091)	0.098	(0.041)	0.099	(0.033)
Current neighbor	0.053	(0.013)	0.089	(0.010)	0.346	(0.079)	0.065	(0.008)	0.111	(0.032)	0.078	(0.021)	0.113	(0.022)
Constant	0.013	(0.001)	0.031	(0.000)	0.160	(0.002)	0.015	(0.000)	0.023	(0.001)	0.020	(0.001)	0.020	(0.001)
No of fixed-effects	501,117		1,586,443		225,541		1,862,019		196,128		328,600		409,768	
No of observations	10,161,663		30,950,111		3,260,438		37,851,336		4,750,956		7,793,215		9,414,202	

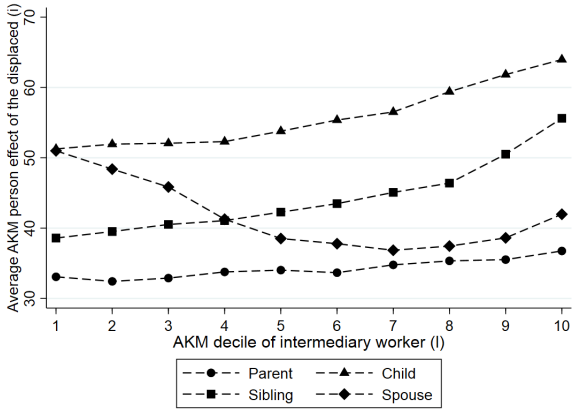
Notes: All estimates are expressed in percentage points, i.e., the coefficients are multiplied by 100. All estimations include establishment-pair(-and-year) fixed-effects. Standard errors are clustered on the potential hiring establishment-and-year level.

B.2 Additional results on sorting of social connections (Section 4.1)

(a) All connections (as 1st col. Table 4, Panel A)



(b) Family members



(c) Former classmates

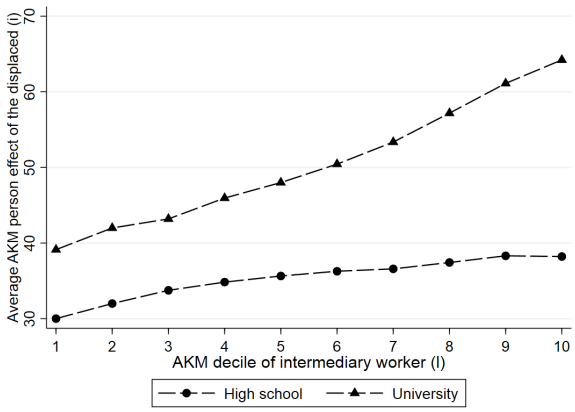
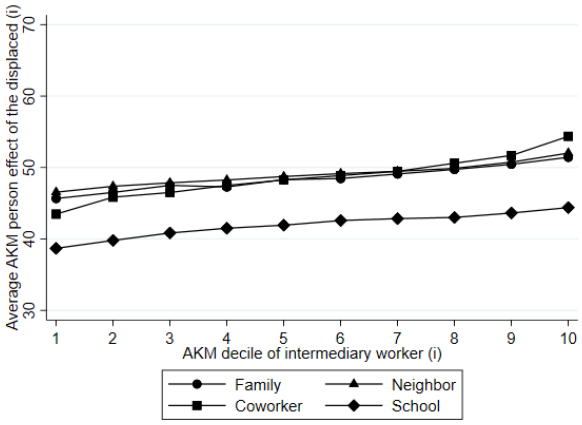
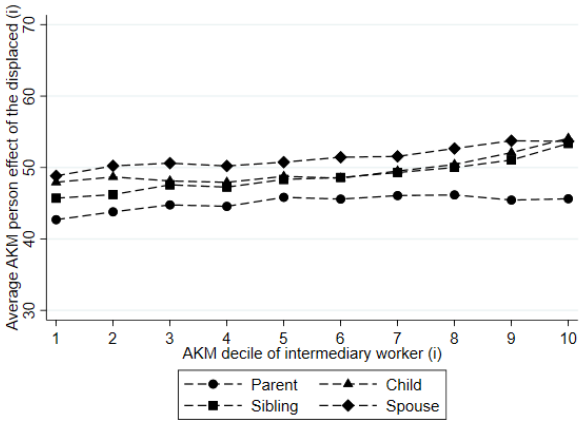


Figure B.1: Mean person effects (θ_i) percentile of displaced workers (y-axis) by decile of person effects ($\theta_l \mid C_{il} = 1$) of connected intermediary workers.

(a) All connections (as 2nd col. Table 4, Panel A)



(b) Family members



(c) Former classmates

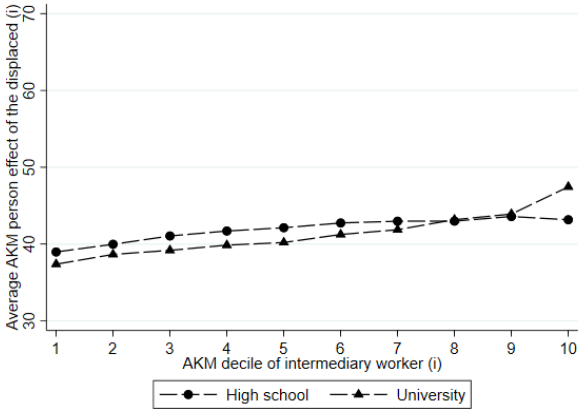
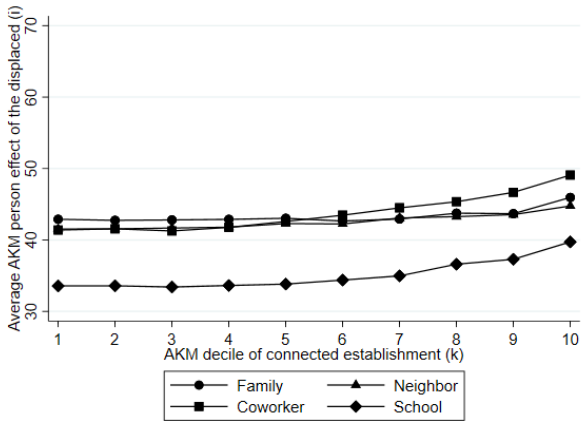
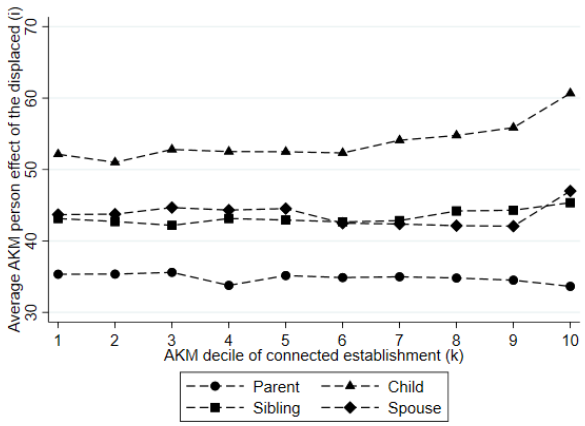


Figure B.2: Mean residualized person effects ($\hat{\theta}_i$) percentile of displaced workers (y-axis) by decile of residualized person effects ($\hat{\theta}_l \mid C_{il} = 1$) of connected intermediary workers.

(a) All connections (as 3rd col. Table 4, Panel A)



(b) Family members



(c) Former classmates

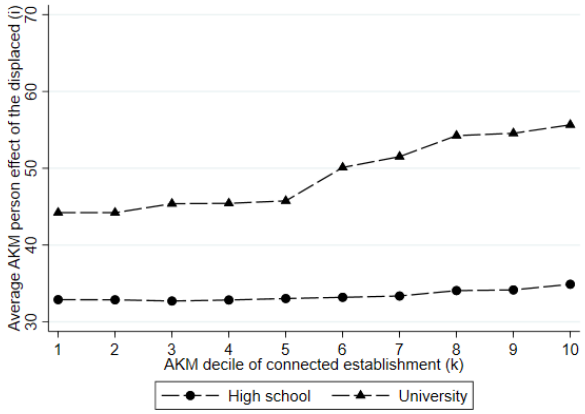
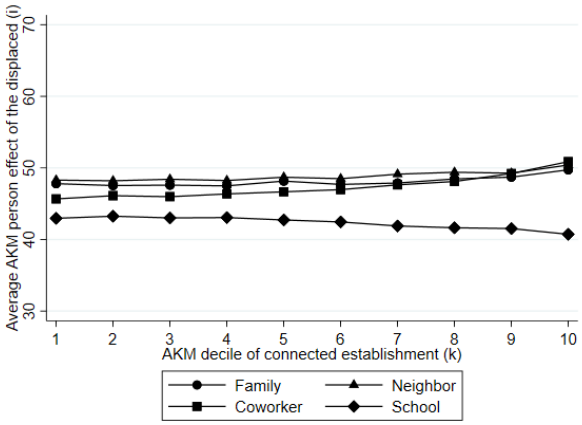
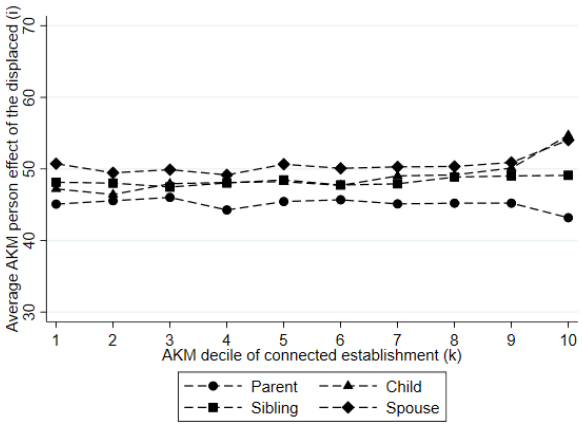


Figure B.3: Mean person effects (θ_i) percentile of displaced workers (y-axis) by decile of establishments effects of connected establishments ($\psi_k \mid C_{ik} = 1$).

(a) All connections (as 4th col. Table 4, Panel A)



(b) Family members



(c) Former classmates

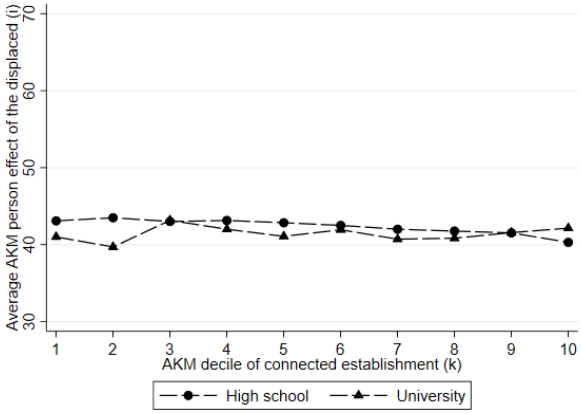
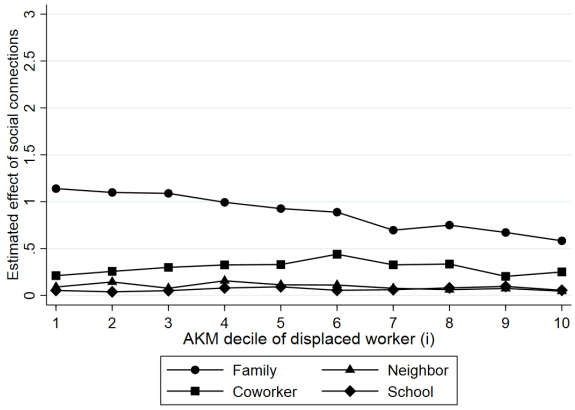


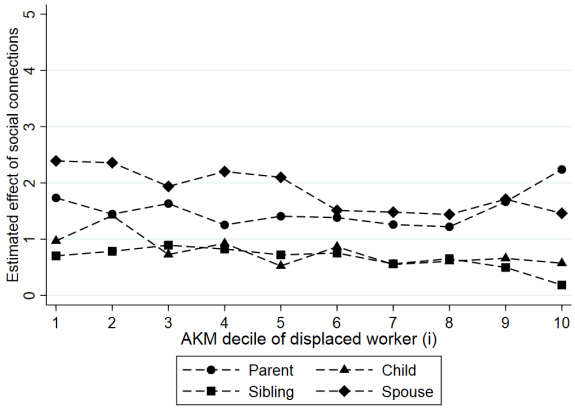
Figure B.4: Mean residualized person effects ($\hat{\theta}_i$) percentile of displaced workers (y-axis) by decile of establishments effects ($\psi_k \mid C_{ik} = 1$) of connected establishments.

B.3 Effects of connections by person and establishment effects and type of connections (related to Figures 4 and 6)

(a) All connections



(b) Family members



(c) Former classmates

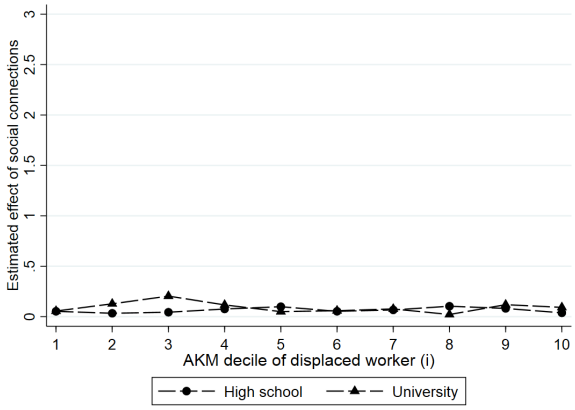
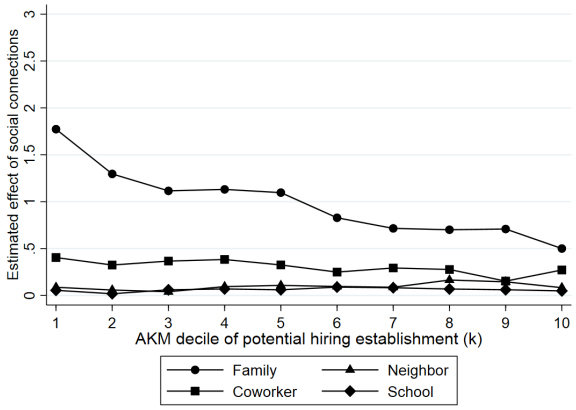


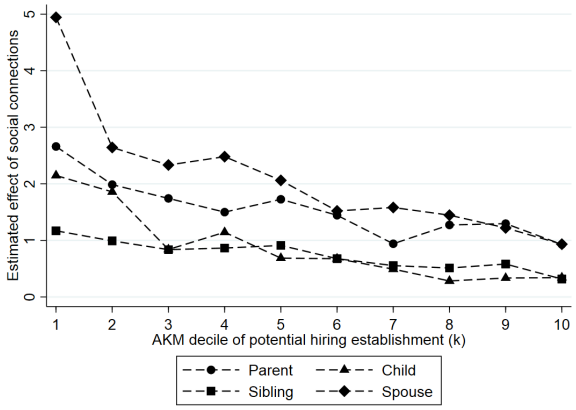
Figure B.5: Estimated effect of social connections by AKM effect (θ_i) decile of displaced worker i.

Note: For coefficients and standard errors of the linear slopes, see Table B.3.

(a) All connections



(b) Family members



(c) Former classmates

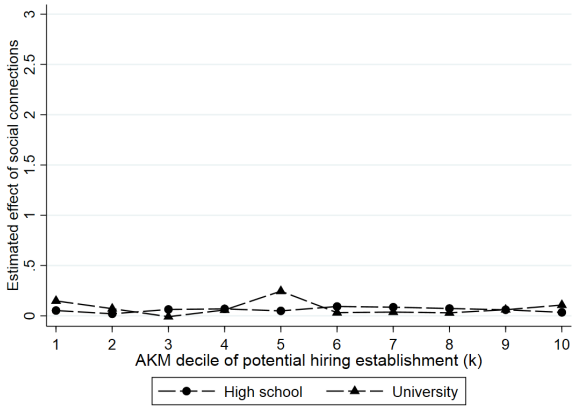


Figure B.6: Estimated effect of social connections by AKM effect (ψ_k) deciles of potential hiring establishment k.

Note: For coefficients and standard errors of the linear slopes, see Table B.3.

B.4 Causal effects by type and person effect (as in Figures 5 and B.5, but after residualizing from demographics).

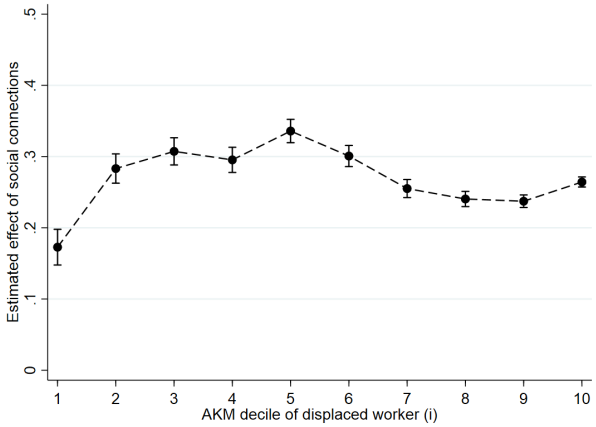


Figure B.7: Estimated effects of social connections, by residualized AKM person effects deciles

Notes: The figure repeats Figure 5 with the one difference that the AKM person effects of the displaced worker (i) have been residualized from age at displacement, education level and gender.

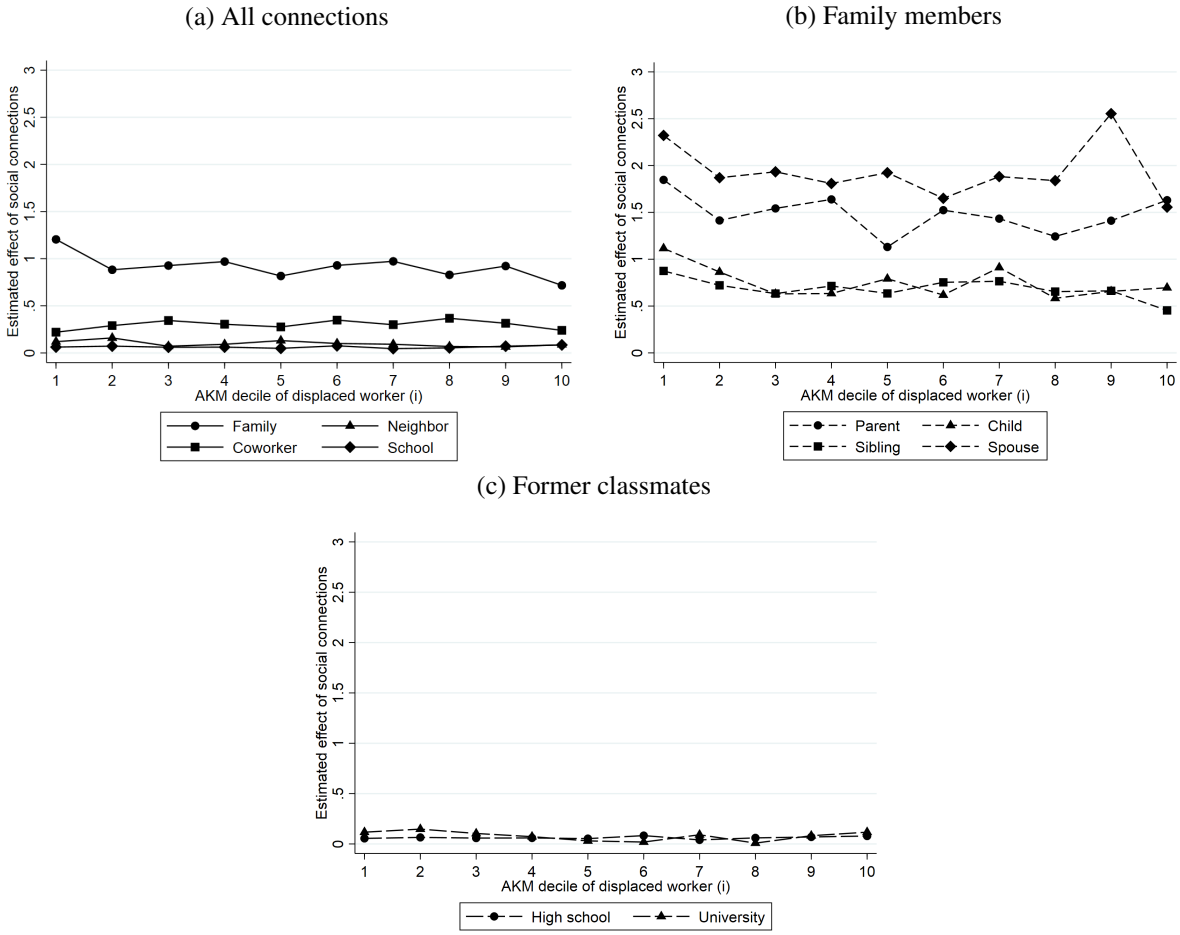


Figure B.8: Estimated effect of social connections by residualized AKM effect ($\hat{\theta}_i$) decile of displaced worker (i)

Notes: The figure repeats Figure B.5 with the one difference that the AKM person effects of the displaced worker (i) have been residualized from age at displacement, education level and gender. For coefficients and standard errors of the linear slopes, see Table B.3.

B.5 Estimates related to results presented in graphs

Table B.3: Linear slope coefficients for Figures B.5, B.8, 4 and B.6

	Figure B.5		Figure B.8		Figure 4/B.6	
	Coef.	(s.e)	Coef.	(s.e)	Coef.	(s.e)
Panel A: All connections:						
Family member	-0.065	(0.008)	-0.020	(0.008)	-0.131	(0.008)
Former co-worker	0.006	(0.004)	0.004	(0.004)	-0.018	(0.003)
Former classmate	0.004	(0.002)	-0.000	(0.002)	0.002	(0.001)
Current neighbor	-0.006	(0.004)	-0.004	(0.003)	-0.000	(0.003)
Panel B: Family members:						
Parent	-0.042	(0.028)	-0.038	(0.025)	-0.182	(0.020)
Adult child	-0.057	(0.022)	-0.006	(0.021)	-0.173	(0.022)
Spouse	-0.116	(0.031)	-0.008	(0.033)	-0.366	(0.033)
Sibling	-0.044	(0.009)	-0.017	(0.010)	-0.083	(0.009)
Panel C: Former classmates:						
High school	0.006	(0.002)	-0.000	(0.002)	0.002	(0.001)
College/university	-0.006	(0.008)	-0.002	(0.006)	-0.003	(0.007)

Notes: Columns (1) and (2) display the estimates for the linear interaction between social connections and AKM person effect deciles. Columns (3) and (4) display the same estimates when the AKM person effects of the displaced worker (i) have been residualized from age at displacement, sex, and education level. Columns (5) and (6) display the estimates for the linear interaction between social connections and AKM establishment effect deciles. Graphical representations of the results (from more flexible specifications using interactions with decile dummies) are shown in Figures B.5, B.8, and B.6. All estimates are expressed in percentage points, i.e., the coefficients are multiplied by 100. Standard errors are clustered on the potential hiring establishment-and-year level.

Table B.4: Estimates underlying Figures 7 and 9

	Figure 7 [a]		Figure 9 [b]	
	Coef.	(s.e.)	Coef.	(s.e.)
Any connection	0.361	(0.029)		
Any connection × (Person effect/100)	0.194	(0.087)		
Any connection × (Person effect/100)^2	-0.189	(0.083)		
Any connection × (Person effect/100) × (Establishment effect/100)	-0.009	(0.077)	0.044	(0.066)
Any connection × (Establishment effect/100)	-0.251	(0.088)	-0.176	(0.079)
Any connection × (Establishment effect/100)^2	-0.029	(0.076)	0.065	(0.071)
(Person effect/100)	0.021	(0.007)		
(Person effect/100)^2	-0.026	(0.008)		
(Person effect/100) × (Establishment effect/100)	-0.004	(0.006)		
Constant	0.032	(0.001)	0.429	(0.017)
No of observations	24,600,167		1,682,201	
R-squared	0.228		0.203	

Notes: All estimates are expressed in percentage points, i.e., the coefficients are multiplied by 100. Standard errors are clustered on the potential hiring establishment-and-year level.

[a] The column displays the estimates from equation (4), where an indicator for social connection is interacted with a second order polynomial of both AKM person effects of the displaced and establishment effects of potential hiring establishments (in percentiles). Graphical representation of the results are shown in Figure 7 in the main text.

[b] The column displays the estimates from equation (6), which include worker and set of connected establishments and year fixed-effects. It shows the interactions with a second order polynomial of the AKM effect of the potential hiring establishment and the interaction between the AKM person effect of the displaced and the establishment effect of the potential hiring establishment (in percentiles). Note that the baseline person effect of the displaced is absorbed by the (worker) fixed-effect.

B.6 Sorting patterns when including post-hire data in the estimation of AKM person effects (related to Figure 10)

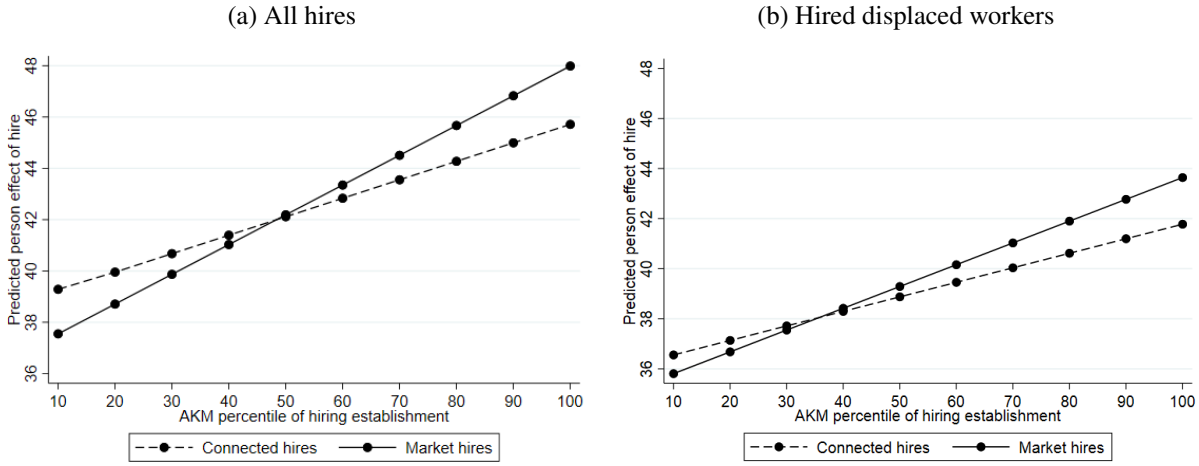


Figure B.9: Predicted person effect of new hire as a function of hiring establishment effects and the use of social connections

Notes: The figure is the same as Figure 10 with the only difference that the person effects estimated in the pre-displacement period have been replaced with person effects using the entire data. We have added the mean person effect within each sample.

B.7 Results adding more detailed connection types than in the paper

Table B.5: Tests based on other non-connected potential hiring establishments (related to Table 6)

	Baseline [a]		...same firm [b]		...same industry [c]	
	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)
Panel A:						
Any connection	0.270	(0.005)	0.029	(0.004)	0.014	(0.001)
Constant	0.026	(0.000)	0.006	(0.001)	0.006	(0.000)
Panel B:						
Family member	1.095	(0.020)	0.037	(0.011)	0.016	(0.004)
Former co-worker	0.253	(0.010)	0.075	(0.012)	0.016	(0.003)
Former classmate	0.066	(0.004)	0.011	(0.004)	0.010	(0.002)
Current neighbor	0.086	(0.008)	0.015	(0.009)	0.023	(0.006)
Constant	0.027	(0.000)	0.006	(0.001)	0.006	(0.000)
Panel C:						
Family members						
Parent	1.867	(0.052)	0.061	(0.024)	0.017	(0.011)
Adult child	0.670	(0.051)	0.027	(0.040)	-0.011	(0.011)
Spouse	1.974	(0.078)	0.080	(0.043)	0.020	(0.012)
Sibling	0.697	(0.023)	0.018	(0.014)	0.019	(0.006)
Former co-worker	0.252	(0.010)	0.075	(0.012)	0.016	(0.003)
Former classmate						
High school	0.064	(0.004)	0.011	(0.004)	0.010	(0.002)
College/university	0.088	(0.018)	0.006	(0.013)	0.004	(0.005)
Current neighbor	0.080	(0.008)	0.015	(0.009)	0.023	(0.006)
Constant	0.026	(0.000)	0.006	(0.001)	0.006	(0.000)
No of fixed-effects		2,087,560		391,438		1,540,941
No of observations		41,111,774		3,676,175		29,891,982

Notes: Data are in dyadic form with one observation per combination of displaced worker and potential hiring establishment. All estimates are expressed in percentage points, i.e., the coefficients are multiplied by 100. All estimations include establishment-pair(-and-year) fixed-effects. Standard errors are clustered on the potential hiring establishment-and-year level.
[a] Repeats the first column of Table **??**.
[b] Each potential hiring establishment has been replaced by another randomly selected establishment within the same firm, location (i.e., municipality), and industry (i.e., 3-digit code).
[c] Each potential hiring establishment has been replaced by another randomly selected establishment within the same location (i.e., municipality) and industry (i.e., 3-digit code).

Table B.6: Post-hire outcomes (related to Table 9)

	Outcomes after 1 year		Outcomes after 3 years					
	Log(Earnings) [a]		Log(Earnings) [a]		Employed [b]		Job stability [c]	
	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)	Coef.	(s.e.)
Panel A								
Any connection	0.083	(0.013)	0.175	(0.019)	0.029	(0.004)	0.128	(0.007)
Constant	11.043	(0.026)	10.805	(0.027)	0.747	(0.005)	0.249	(0.005)
Panel B								
Family members	0.002	(0.018)	0.152	(0.027)	0.028	(0.006)	0.136	(0.008)
Former co-worker	0.137	(0.019)	0.148	(0.031)	0.020	(0.006)	0.076	(0.011)
Former classmate	0.179	(0.029)	0.201	(0.049)	0.021	(0.010)	0.158	(0.017)
Current neighbor	0.126	(0.037)	0.197	(0.059)	0.036	(0.013)	0.152	(0.023)
Constant	11.043	(0.026)	10.805	(0.027)	0.747	(0.005)	0.249	(0.005)
Panel C								
Family members								
Parent	-0.123	(0.030)	0.073	(0.043)	0.020	(0.009)	0.101	(0.014)
Adult child	0.055	(0.059)	0.177	(0.096)	0.029	(0.019)	0.109	(0.029)
Spouse	0.086	(0.034)	0.104	(0.058)	0.031	(0.011)	0.117	(0.018)
Sibling	0.042	(0.027)	0.161	(0.045)	0.025	(0.009)	0.135	(0.015)
Former co-worker	0.135	(0.019)	0.149	(0.031)	0.020	(0.006)	0.077	(0.011)
Former classmate								
High school	0.180	(0.030)	0.217	(0.053)	0.022	(0.011)	0.158	(0.018)
College/university	0.156	(0.092)	-0.020	(0.124)	-0.008	(0.019)	0.137	(0.044)
Current neighbor	0.132	(0.037)	0.208	(0.058)	0.037	(0.013)	0.159	(0.023)
Constant	11.043	(0.026)	10.806	(0.027)	0.747	(0.005)	0.250	(0.005)
No of fixed-effects		29,554		29,554		29,554		29,554
No of observations		208,738		208,738		208,738		208,738

Notes: The estimation sample is all displaced workers who were employed in in November of year $t+1$. All estimations include closing establishment (-and-year) fixed-effects and controls for the workers' age, sex, education, and three years of pre-displacement employment, earnings, and employer history. Standard errors are clustered on the closing establishment (-and-year) level.

[a] Earnings is defined as annual labor income and has been left censored at SEK 1,000.

[b] Employed is defined as being employed in November of year $t+3$.

[c] Job stability is defined as being employed at the same establishment in November of both year $t+1$ and $t+3$.

Appendix C Establishment-level characteristics

One of our main findings is that the causal impact of social connections varies dramatically with the wage-level of the connected establishment. Therefore, we have also repeated the main analysis for various subsamples based on other characteristics of the potential hiring establishment. In Figure C.1, we present estimates for an average across all connections. The analysis for each specific type of connection can be found in Appendix Table B.2.

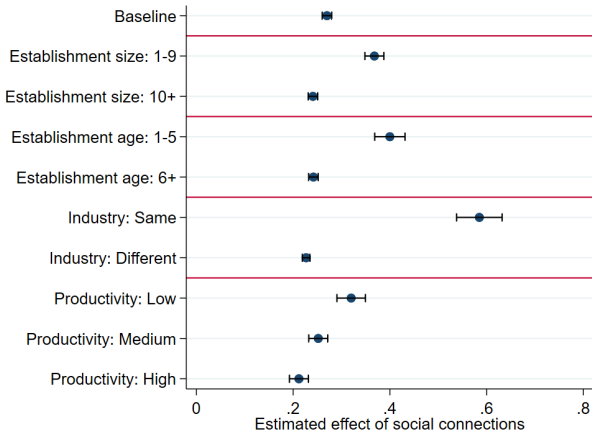


Figure C.1: The estimated importance of social connections by characteristics of the potential hiring establishment (k)

Notes: All estimates are expressed in percentage points, i.e., the coefficients are multiplied by 100. All estimations include establishment-pair(-and-year) fixed-effects. Standard errors are clustered on the potential hiring establishment-and-year level. See the Appendix for results disaggregated by connection type.

The figure first shows that the effects are more potent for small establishments and for comparably young establishments. A reason may well be that there are establishments where hiring frictions are more pronounced.

Next, we split the sample into cases where closing establishment j and potential hiring establishment k are in the same industry, vs when they operate in different industries. This sample-split is clearly related to the potential fear that the closure of an establishment may affect product market competition in the industry and location as discussed in Section 4 above. Separating between inter- and intra-industry connections serves as an additional test on top of the "placebo-style" regressions: if competition effects were indeed a major force, our results should mostly come from within-industry connections. However, when the closing and potential hiring establishment operate within the same industry, the displaced workers' skill sets obviously satisfy the requirements of the potential hiring establishment.[59] Results in Figure C.1 show that the estimated importance of connections is much larger (0.58) for within-industry

[59]Not surprisingly, the baseline probability of intra-industry hiring (0.162) of non-connected displaced workers is more than 10 times the equivalent probability of inter-industry hiring (0.015).

connections than for between-industry connections (0.23). However, the impact of between-industry connections is clearly significant and, even more important, the between-industry estimate is very close to the baseline impact. The reason is that most connections span across industries.

Finally, we show how the use of connections in the hiring process varies with firm's productivity.[60] It is well-known that productivity measures tend to correlate heavily with firm or establishment effects from AKM-decompositions. Thus, we expect the effects to be larger in cases where the productivity of firm k is lower than average. This conjecture is confirmed in Figure C.1 which shows that the causal impact of connections is negatively related to the productivity of the connected firm.

[60]The productivity measure is only available for a sub-sample of our firms. It is measured at the firm-level rather than the establishment-level. Productivity is then categorized (i.e., low, medium, and high productivity) based on the firm's position in the distribution of value-added per worker within the local labor market and industry.

Appendix D The displaced and the intermediary worker's characteristics and their similarity

Here we study the extent to which the impact of social connections varies with observable characteristics in the following dimensions: (i) the connected displaced worker i's own characteristics, (ii) the intermediary worker l's characteristics, and (iii) their similarity in terms of these characteristics. We re-estimate our model including interactions between the indicator for having a connection (of any kind) and indicators for various characteristics (i.e., sex, age, nativity, and attained education level) of the two workers constituting the connection (i.e., the displaced worker i and the intermediary worker l), and their similarity in terms of these characteristics.

For simplicity, we drop establishment-pairs where the potential recruiting establishment is connected to a displaced worker through multiple intermediary workers. We focus on connections through former co-workers since this is the only type of connection that does not impose strong boundaries on the characteristics of the two workers and their similarity. For example, a parent-child connection has strong implications for the ages of the two, and formerly being classmates has strong implications for both the age and level of attained education. We present the results of the analyses of the mediating role of the displaced and intermediary workers' characteristics (and their similarity) in Table **??**.

Before turning to the mediating role of displaced and intermediary workers' characteristics, it may be useful to note that the inclusion of worker characteristics has no impact at all on our estimates of interest. This supports our choice to rely only on establishment-pair(-and-year) fixed-effects in previous estimations.

Turning to the mediating role of the displaced workers' characteristics, the impact of a co-worker connection seems to be mediated the most by being female (negatively) and being aged 50–64 years (negatively). These results support the findings in previous literature (Bentolila et al., 2010; Ioannides and Datcher Loury, 2004).

Shifting the focus to the intermediary workers' characteristics, we find that if the intermediary worker is a man, prime aged, or has no more than compulsory schooling the connected displaced worker seems to be more likely to get hired. These results are in line with the (in this case, very scarce) previous literature provided by Kramarz and Skans (2014) and Bayer et al. (2008).

There is a long-standing sociological notion (McPherson et al., 2001) that similarity in all dimensions reinforces the importance of social interactions. Our estimates of how the impact of social connections varies with similarity, between the displaced and the intermediary worker, in terms of sharing the same characteristics support this notion. The estimates are positive and statistically significant for all characteristics, in particular in the immigration, gender, and age dimensions (more than education). This

further reinforces the consistent result that social proximity, or tie strength, is crucial for the usefulness of social connections. clear