SOCIAL SCIENCES

The Trojan-horse mechanism: How networks reduce gender segregation

M. Arvidsson[1]*, F. Collet[2]*, P. Hedström[1]*

Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

The segregation of labor markets along ethnic and gender lines is socially highly consequential, and the social science literature has long viewed homophily and network-based job recruitments as some of its most crucial drivers. Here, we focus on a previously unidentified mechanism, the Trojan-horse mechanism, which, in contradiction to the main tenet of previous research, suggests that network-based recruitment reduce rather than increase segregation levels. We identify the conditions under which networks are desegregating, and using unique data on all individuals and all workplaces located in the Stockholm region during the years 2000–2017, we find strong empirical evidence for the Trojan-horse mechanism and its role in the gender segregation of labor markets.

INTRODUCTION

Individuals often find their jobs through friends, colleagues, and acquaintances, and the importance of labor market networks for a wide range of socioeconomic outcomes has been documented in numerous studies (*1*–*8*). Analyses of the role of labor-market networks in segregation processes usually center on homophily (*8*–*13*; see *14*, *15* for rare exceptions), and the homophily thesis implies that individuals with similar characteristics agglomerate at certain workplaces/organizations. That is, if an individual with property X joins an organization, then the probability of additional individuals with property X joining the organization increases, and this kind of self-reinforcing process is expected to generate a more segregated market. What this line of research has overlooked, however, is that opportunity structures often trump individual preferences. Focusing on gender segregation and the mixing constraints that employees face in tie formation within organizations, we show that networks created by individuals moving between organizations—which we refer to as mobility networks—under a wide set of circumstances in fact are desegregating. In particular, we identify a mechanism, the Trojan-horse mechanism, that shows how networks counteract the impact of segregating mobility events. We refer to this mechanism as the Trojan-horse mechanism because it shows, just as in the case of the soldiers in the hollow Trojan horse, how the recruitment of a new employee belonging to one specific category (e.g., a male) can open up the gates for those of another category (e.g., females) and thereby change the composition of the organization in unexpected ways. Our analyses reveal that if this mechanism is commonly observed in a market, network-based mobility is likely to desegregate rather than segregate the market. To test this prediction, we use a large-scale longitudinal register dataset with rich demographic and socioeconomic information, as well as detailed mobility records, for every individual and every organization that resided in the greater Stockholm Metropolitan area during the years 2000–2017.

Trojan horses and segregation processes

The main reason for mobility networks having a desegregating effect is because of the mixing constraints that employees face within organizations, and particularly, what we refer to as the Trojan-horse mechanism, a sequence of interlinked events through which the mobility of a minority group member triggers subsequent moves along the same network path by members of the majority group.

The seminal work by Peter Blau and colleagues [e.g., (*16*, *17*)] provides the key for understanding why this is likely to happen. In a series of publications, they studied the interplay between homophily and group structure in the formation of cross-category ties. Their analyses showed that although individuals often prefer to form ties to similar others, the composition of the groups they are part of often prevents them from doing so. Their core prediction was that if individuals are in a minority within their groups, then the possibility and probability of them forming homophilous ties are reduced and their probability of forming cross-category ties is increased. This prediction has been supported in numerous empirical studies [e.g., (*18*, *19*)]. In the context of labor markets, Blau's research suggests that if an individual of category X works in an organization in which most individuals belong to category Y, then the individual in question is likely to form more ties to Y-individuals than to X-individuals even if he/she prefers to have ties to X-individuals.

The Trojan-horse mechanism extends Blau's theory to the case of intergroup dependencies created by individuals' mobility between groups. Mobility can be represented as a network where the nodes are organizations and ties are formed when individuals move from one organization to another. Prior research (*6*, *7*) suggests that privileged job-related information can flow through these links and that this information is likely to influence subsequent mobility events (*1*). More specifically, if we let j and k represent two different organizations, prior research suggests (i) that the probability of relevant information about k reaching individuals in j increases if there are individuals in k who previously worked in j, (ii) that this information is likely to affect subsequent mobility from j to k, and (iii) that individuals in j who are similar to the individual who moved from j to k in terms of sociodemographic characteristics are particularly likely to be informed about and affected by what goes on in k.

Together, these two insights—that group composition is predictive of the formation of cross-category ties and that prior mobility promotes future mobility along the same network paths—result in the Trojan-horse prediction: When individuals leave an organization in which they are in a minority, they are more likely to be followed by majority-group individuals and this, in turn, implies that initially segregating moves can set in motion a chain of desegregating moves.

[1]The Institute for Analytical Sociology, Department of Management and Engineering, Linkoping University, SE-601 74 Norrköping, Sweden. [2]People Management and Organisation Department, ESADE, Ramon Llull University, 08034 Barcelona, Spain.
*Corresponding author. Email: martin.arvidsson@liu.se (M.A.); francois.collet@esade.edu (F.C.); peter.hedstrom@liu.se (P.H.)

Downloaded from https://www.science.org on February 26, 2025

The distinction between segregating and desegregating moves can be precisely defined using James and Taeuber's (*20*) so-called principle of transfer: Everything else being equal, the labor market becomes more sociodemographically segregated if an individual moves to an organization with a higher proportion of same-category employees than in the organization he/she left (a segregating move), and the market becomes less segregated if the organization they move to contains a lower proportion of same-category employees (a desegregating move).

A stylized example is useful for illustrating how the Trojan mechanism works (see Fig. 1). If a woman moved from organization *j* to *k* at time *t*, then this would make the labor market more gender-segregated because the destination (*k*) contains a larger percentage of women than the origin (*j*).

However, a move like this should not be considered in isolation, since it is likely to affect the moves of other individuals. In this particular example, and as suggested by Blau's research, even if the woman who moved from *j* to *k* preferred ties to other women, the gender composition of organization *j* was such that she probably had more ties to men than to women. Therefore, to the extent that someone from *j* follows in her path, this person is likely to be a man, and a move by a man is desegregating because *k* contains a smaller percentage of men than *j*. Once a male tie is formed from *j* to *k*, subsequent moves are even more likely to be male because of the large number of men in *j* and because of traditional homophilous tie-creation tendencies (*21*). Hence, this example illustrates how the Trojan mechanism makes the market less segregated: An initial segregating move between two organizations opens up a mobility path for those of the opposite gender and sets in motion a desegregating process.

RESULTS

Because the Trojan-horse mechanism operates at one level (micro) while its ultimate outcomes are observed at another level (macro), our analyses are divided into two parts. First, we subject the core hypothesized effects of the Trojan mechanism to detailed empirical tests using a node embedding–based dynamic matched sample design (*22*, *23*). Second, we examine the hypothesized macro-level implications of the Trojan mechanism using a large-scale, empirically calibrated simulation model that allows us to simulate counterfactual "worlds" (*24*) where everything is kept constant except for the networks in which the individuals are embedded.

To perform these analyses, fine-grained individual-level data at a population scale is required. We use a unique database that includes rich demographic and socioeconomic information on every individual and every organization that ever resided in the greater Stockholm Metropolitan area during the years 2000–2017. Statistics Sweden assembled the database for us by merging a large number of administrative and population registers. The number of organizations included in the analyses ranges from 20,000 to 30,000 each year, and the number of individuals is about 700,000 at any point in time (see Supplementary Text for details).

Testing the micro-level assumptions

There are two hypothesized effects at the heart of the Trojan mechanism: (i) that a move between two organizations increases the probability of additional moves along the same path and (ii) that the group composition of the organization of origin is predictive of the sociodemographic characteristics of the followers.

To test the first hypothesized effect, we use a node embedding–based dynamic matched sample design (see Materials and Methods) to contrast the likelihood of a move between two organizations with a prior mover (the treated) with the likelihood of a move between two organizations without a prior mover (the control). Corroborating the hypothesized effect, Fig. 2A shows that the presence of a prior mover substantially increases the probability of future moves along the same network paths. The estimated treatment effect, measured in terms of relative risk, is 3.7 [± 0.2 (95% confidence interval)] for those of the same gender as the initial mover and 3.3 (±0.2) for those of the opposite gender, suggesting that if a colleague moved to a particular organization during the previous year, then the probability that at least one individual of the same-gender moves to the same organization in the next year increases by a factor of 3.7 and at least one of the opposite gender by a factor of 3.3. Contrasting these estimates to those obtained by random matching, Fig. 2A also shows that, without adjusting for homophilous mobility tendencies, the estimated treatment effect would be grossly overestimated (increasing by more than 20-fold).

To test the second hypothesized effect, we examine whether the group composition of the organization of origin is predictive of the sociodemographic characteristics of those who follow. We do this by examining the gender of the followers for all dyadic pairs with a prior mover (the treated) during the years 2000–2017. As shown in Fig. 2B, the gender of those who follow an individual's path is indeed highly contingent upon the gender composition of the organization of origin, and the patterns are in line with the theoretical expectations.

To make the observed data patterns more transparent, Fig. 2B distinguishes between three types of *j* (origin) organizations: organizations with (a) less than 10%, (b) between 40 and 60%, and (c)

Downloaded from https://www.science.org on February 26, 2025

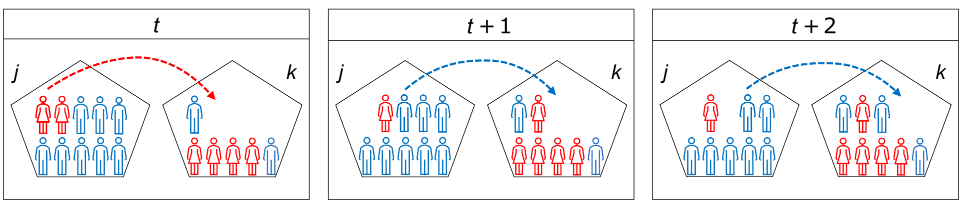

Fig. 1. Schematic representation of the Trojan-horse mechanism. (***t***) With no prior mobility between organization *j* and *k*, a female moves from *j* to *k* (segregating move), increasing the probability for future mobility along the same path. (***t* + 1**) Strong mixing constraints (8 of 9 males in *j*) makes it likely that the female that moved had more ties to males than females, despite homophily. Therefore, it is more likely that a male follows in her path (desegregating move). (***t* + 2**) The combined effect of composition (7 of 8 males in *j*) and preference for same-gender ties makes it much more likely that another male will follow (desegregation move). While this example shows a stylized Trojan sequence of length three, it is important to note that, as the counteracting (desegregating) effect occurs already at the second step, a sequence of length-two also constitutes a realization of the Trojan-horse mechanism.

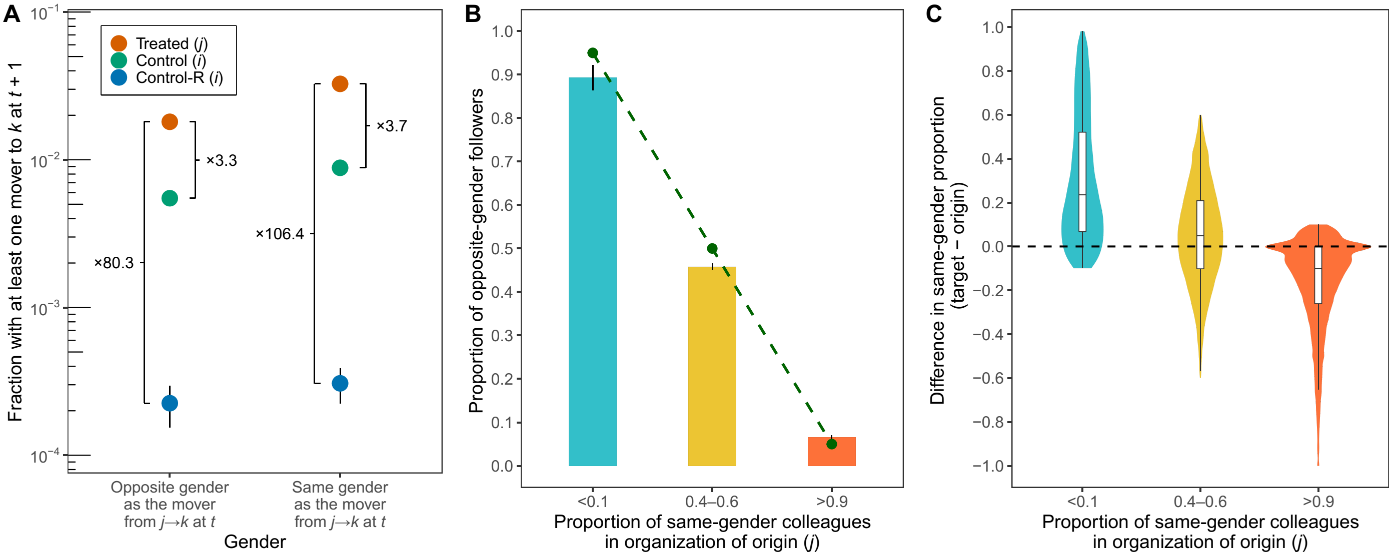

Fig. 2. Empirical evidence of the Trojan-horse mechanism. (**A**) Fraction of the treated pairs $j \rightarrow k$ (with a prior move) and fraction of the control pairs $i \rightarrow k$ (without any prior move) where at least one individual moved to the destination (k) at time $t + 1$ (treatment occurs at time t). Error bars represent 95% confidence intervals obtained from estimated logistic regression models (see Supplementary Text for details). The x axis distinguishes between same- and opposite-gender movers (relative to the gender of the prior mover from j to k). The numbers shown alongside the vertical lines represent the estimated treatment effects, calculated as the risk ratio between the treated and the control group (i.e., by comparing the fraction of cases with at least one mover at $t + 1$ under each treatment regime). Control-R corresponds to random matching, and Control corresponds to coarsened exact matching. (**B**) Proportion of opposite-gender followers, relative to the initial mover, for all treated $j \rightarrow k$ dyads, plotted as a function of the same gender proportion (SGP) in the organization of origin (j). Three ranges of SGP are considered: less than 0.1, between 0.4 and 0.6, and more than 0.9. Error bars represent 95% binomial confidence intervals calculated using the normal approximation. (**C**) Distribution over differences in SGP between the target organization (k) and the organization of origin (j) for initiating moves, i.e., those that create new ties, for the same three ranges of SGP as in (B).

more than 90% of the same gender as the individual who moved from j to k during year t. The green reference line indicates the expected proportion of opposite-gender followers for each of these three categories if mobility perfectly mirrored the gender composition of the organization of origin. For example, if the first mover was a woman and the proportion of women in the organization of origin was 0.05, then the expected proportion of opposite gender followers is 0.95.

When an individual leaves an organization where he/she was in a small minority (type a), it is very likely that someone of the opposite gender follows in his/her path. As shown in Fig. 2B, about 89.5% of the followers then are of the opposite gender. If the gender composition of the organization is balanced (type b), then the probability of an opposite-gender follower is much lower. About 47.5% of the followers then are of the opposite gender. Last, when an individual leaves an organization with a strong same-gender majority (type c), it is very unlikely that someone of the opposite gender will follow in his/her path. Less than 10% of the followers then will be of the opposite gender. Overall, the evidence shows that the gender of followers is strongly influenced by the gender composition of the organization of origin.

Local effect on segregation

Whether a Trojan-sequence has a desegregating effect or not crucially depends on the gender compositions of the organizations involved. If the initial move is segregating, opposite-gender followers will desegregate the market, and vice versa. Figure 2C displays the empirically observed gender-composition differences (destination-origin) for such initiating moves using the same categorization of organizations (i.e., type a, b, and c) as in Fig 2B.

The vast majority of individuals who leave an organization in which they are in a small minority [same gender proportion (SGP) < 0.1] tend to move to organizations with a higher percentage of same-gender individuals (see the left violin plot in Fig. 2C). Such moves increase the segregation of the market, but because of their minority status in j, they tend to—as shown in Fig. 2B—be followed by those of the opposite gender, and their moves desegregate the market. As a result, Trojan sequences almost always tend to be desegregating.

In contrast, individuals who leave organizations in which they are in a strong majority position (SGP > 0.9) tend to move to organizations with a lower proportion of same-gender individuals (see Fig 2C), and these moves are desegregating. Those who follow in their path are likely to be of the same gender, thereby further desegregating the market. This pattern shows that how a move to an organization with a lower proportion of same-gender individuals can set in motion network-based processes that amplify the desegregating effect of the first move. Last, if the gender composition of the j organization was balanced (0.4 < SGP < 0.6), Fig 2C shows that moves from such organizations have a relatively limited impact on the segregation of the market.

Macro-level implications

To examine the macro-level implications of the Trojan-horse mechanism we implement a large-scale empirically calibrated simulation model (*25*, *26*). In the simulations, individuals and organizations retain their true characteristics, and they enter and leave the market just as they did in reality. However, individuals do not move between the organizations as they did in reality but as predicted by a conditional logit model with parameters that capture key mobility mechanisms, including the network dependencies of focal interest. More specifically,

Downloaded from https://www.science.org on February 26, 2025

network dependency is captured by two binary variables, same-gender tie (SGT) and opposite-gender tie (OGT), that take the value 1 if a former colleague of the focal individual moved to the potential target organization during the previous time period (see Materials and Methods for details). For each individual-organizational pair, we assign a matching probability based on the conditional logit model, and using multinomial sampling, one destination is selected for each individual. Once all individuals have been assigned their destinations, organizational-level variables and network variables are updated, and the same steps are repeated for the next year until the end of the observation window in 2017 (see Materials and Methods for details).

Before engaging in any counterfactual simulation, we validate the simulation model by confirming that it generates mobility patterns that closely approximate our empirical observations. As shown in fig. S3, we find a high level of consistency both in terms of the generated mobility at the microlevel and with regard to the macro-level demographic composition of the market. Furthermore, and crucially, we find that the simulated mobility patterns capture the empirically observed Trojan-horse dynamics displayed in Fig. 2 well (fig. S3).

To assess the macro-level implications of the Trojan-horse mechanism, we implement a set of network interventions that target specific paths of the network and make them more/less consequential for mobility. Specifically, we alter the impact of the network links that are created when someone makes a segregating move from a minority-gender position (SGP < 40%). We refer to these links as "Trojan-initiating links" because the Trojan mechanism predicts that such links are likely to set in motion desegregating processes since individuals of the opposite gender then are likely to follow in the same paths. Thus, by altering the Trojan-initiating links, we can assess the impact that the Trojan-horse mechanism has on the gender segregation of the labor market. We do this by simulating the following counterfactual worlds:

- World 1 (W1) represents the simulated equivalent of the real world in the sense that the matching of individuals and organizations is assumed to be governed by the actual conditional logit estimates.

- W2 and W3 are identical to W1 except that the importance of the Trojan-horse mechanism is amplified by multiplying the network parameters for Trojan-initiating links by two and three, respectively.

- W4 is identical to W1 except that the Trojan-horse mechanism it blocked by multiplying the OGT network parameters for Trojan-initiating links by 0.

Figure 3A shows that amplifying the effect of the Trojan-horse mechanism (W2 and W3) considerably reduces the segregation levels in the Stockholm labor market, and the greater the amplification, the less segregated the market becomes. Conversely, blocking the Trojan-horse mechanism (W4) results in a substantially increased segregation of the labor market.

Figure 3C presents further results of each intervention: Consistent with the Trojan-horse prediction, in W2 and W3, most (68.3 and 72.5%) of the followers along the intervened paths were of the opposite-gender as the first (segregating) mover, and the amplification increased their frequency by a ratio of 2.0 in W2 and 2.6 in W3, explaining why the segregation levels decreased. In contrast, when the Trojan-horse mechanism is blocked (W4), the proportion of opposite-gender followers drops markedly, down to 11.8%, implying that the initial segregating moves instead triggered more moves of those of the same gender as the initial mover, and thereby the segregation levels increased. Together, these analyses offer strong support for the Trojan-horse prediction and show that the mechanism matters greatly for the overall segregation of the market.

To investigate the overall effect of the networks and the role that the Trojan-horse mechanism plays in this, we simulate four additional counterfactual worlds where the importance of all network paths is increased and where the Trojan-horse mechanism either is activated or blocked:

- W5 and W6 are identical to W1 except that the importance of the network is increased by multiplying the network parameters by 3 and 5, respectively.

- W7 and W8 are identical to W5 and W6, respectively, except that the Trojan-horse mechanism is blocked by setting the opposite-gender network parameters to 0 for Trojan-initiating links.

Downloaded from https://www.science.org on February 26, 2025

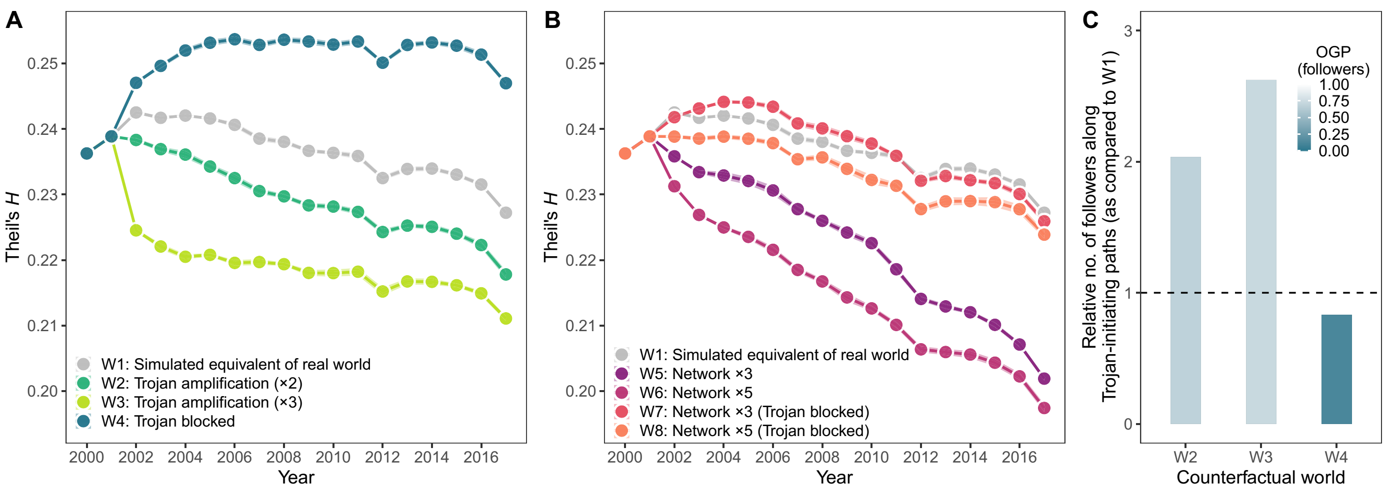

Fig. 3. Simulated trajectories of gender segregation as measured by the entropy index (Theil's *H*) on the Stockholm labor market between 2000 and 2017. Error bands represent 95% confidence intervals based on 10 repeated runs. (**A**) Counterfactual trajectories when the network coefficients for Trojan-initiating links are either increased (W2 and W3) or set at to 0 for opposite-gender colleagues (W4). (**B**) Counterfactual gender segregation level trajectories when the importance of the whole network is increased and the Trojan-horse mechanism is either operative (W5 and W6) or blocked (W7 and W8). (**C**) Relative number of followers along Trojan-initiating links for W2 to W4 compared to the baseline simulation, W1. Colors indicate the proportion of the followers who were of the opposite gender of the initial mover.

Figure 3B shows that, when the Trojan-horse mechanism is active, increasing the importance of the network by a factor of three (W5) or five (W6) leads to substantial reductions in the segregation levels as compared to the baseline simulation (W1). However, if we block the Trojan-horse mechanism, the same interventions—increasing the importance of the network by a factor of three (W7) or five (W8)—only leads to minor changes in the segregation levels as compared to the baseline simulation (W1). These results provide further evidence on the importance of the Trojan-horse mechanism and show that the desegregating effect of networks is crucially dependent on its presence.

DISCUSSION

The findings presented in this study are important for several reasons. First, they provide a nuanced assessment of the role of networks in segregation processes, and the proposed mechanism calls attention to new ways through which networks affect segregation levels. Second, they call into question the main prediction of the existing segregation literature with its emphasis on the importance of self-reinforcing homophily processes. What the previous line of research has overlooked is that opportunity structures often dominate individual motivations. Although individuals may prefer to form ties with individuals who resemble themselves, the groups that the individuals travel within may not allow these motivations to play themselves out. We have shown how this seemingly subtle difference profoundly affects the formation of cross-category ties, the personnel flows between organizations, and the segregation of the market.

We believe that these results are widely generalizable to cases other than gender segregation. Analyses of ethnic segregation, for example, yielded qualitatively very similar results to those reported here (fig. S4). In addition, and more generally, in any setting where we encounter groups (e.g., firms, schools, and neighborhoods) that are composed of individuals with different characteristics (e.g., ethnicity, gender, and social class) and where the individuals move between the groups, the Trojan-horse mechanism is of potential relevance for explaining how individual characteristics become distributed across groups.

While the Trojan horse is a highly general mechanism, a number of factors can affect the generalizability of our results. For example, in settings with highly institutionalized forms of segregation such as in the United States before the Civil Rights Act of 1964 where employees with certain demographic characteristics were barred from certain organizations and jobs, networks are likely to have been rather irrelevant (*27*). Furthermore, different categories of employees may not share job-relevant information to the same extent (*14*, *15*, *28*), and this could alter the dynamics that we observe. These observations suggest that while the Trojan mechanism is highly general and likely to apply to most labor markets, more work is needed to establish how the effect of the Trojan mechanism varies across institutional settings.

Since segregation is of considerable importance for a range of social outcomes such as school outcomes (*29*), neighborhood decline (*30*), and the persistence of gender inequality (*2*, *5*), understanding the mechanisms through which societies become segregated along various socioeconomic and cultural lines is of crucial importance for the design of effective social policies. The nuanced results presented here illustrate how the combination of large-scale longitudinal data with counterfactual/empirically calibrated simulation models can generate new and important insights that would have been difficult to arrive at with smaller survey-based databases paired with the traditional econometric toolkit that so much of the social sciences traditionally has relied upon.

MATERIALS AND METHODS

Testing the micro-level assumptions

When someone follows in the path of another colleague, this could be because of a social tie to the prior mover, but it could also be because they both independently preferred to work in the same organization (*22*, *31*). To distinguish peer effects from alternative mechanisms, we adjust for the key confounding mechanisms that prior research has shown to be important for job mobility (*5*, *32*–*36*) and d node embedding dimensions that serve as proxies for unobserved drivers of mobility (*23*, *37*, *38*). Following the seminal work by Aral *et al.* (*22*), we use a dynamic matched sample design using coarsened exact matching (*39*) for estimation. A pair of organizations $j \rightarrow k$ is considered "treated" at time $t + 1$ if at least one individual of organization j had moved to k during the previous time period, t. This treatment assignment is clearly nonrandom since organizations differ in terms of sectors, industries, and organizational characteristics that strongly predict both the treatment (presence of a prior move to k) and the outcome (subsequent moves to k). Therefore, for each year between 2000 and 2017, we match each treated organizational pair $j \rightarrow k$ with one or more counterfactual organizational pairs $i \rightarrow k$, where i is an organization with the same coarsened-exact organizational profile as j, but which did not experience any prior move to k. To ensure treatment validity, we only consider organizational pairs with no mobility events in the 2 years preceding the treatment year. We match on organizational-level variables that capture mechanisms which previous research has shown to be key drivers of job mobility: (i) homophily (gender, ethnicity, age, education, and salary level) that sort individuals into organizations where many individuals share their profile (*9*, *12*, *40*–*42*), (ii) industry and sector (*42*) that capture noncompositional similarities between organizations, (iii) geographical proximity (*33*), and (iv) the size of organizations and their employee inflows and outflows, which determine baseline mobility probabilities (*32*). Furthermore, we also match on an inferred latent network location for each organization to adjust for unobserved drivers of mobility (*37*, *38*). For example, for a hypothetical move between two financial service firms $j \rightarrow k$ at time t, the counterfactual i would be another financial service firm that is similar to j in size, composition (gender, ethnicity, education, age, and salary), geographical location, latent network location, and with a history of in- and outflow of employees similar to j but without any move from i to k at time t. Details of the matching scheme, the variables used, and balance checks are provided in the Supplementary Text. Last, the effect of prior moves is computed as a relative risk, comparing the fraction of cases where at least one individual moved to the destination in question during the subsequent time point $t + 1$ in organizations with a prior move and in organizations without a prior move. Following established practice (*43*), we obtain SEs from logistic regression models that account for the matched sample structure. Since we expect the treatment effect (i.e., peer effect) to be stronger for individuals of the same gender as the original mover, we estimate separate treatment effects for those of the same gender and for those of the opposite gender of the original mover.

Downloaded from https://www.science.org on February 26, 2025

Testing the macro-level implications

To test the macro-level implications of the Trojan-horse mechanism, we use an empirically calibrated simulation model (*25*, *26*). In the simulation, individuals move between the organizations not as they did in reality but as predicted by a conditional logit model. We define a function, Π_{aijt+1}, that gives the propensity that individual a in organization i is matched with organization j at time $t + 1$

$$\Pi_{aijt+1} = \alpha X_{jt} + \beta X_{it} X_{jt} + \gamma Z_{aijt} + \delta X_{jt} W_{at} + \omega Z_{aijt} X_{jt} + \varphi_{aijt+1}$$

where X_{it} and X_{jt} are organizational properties, Z_{aijt} are two network variables indicating whether there is a SGT and/or an OGT—with respect to individual a—from organization i to j, W_{at} are individual properties, and φ_{aijt+1} is a type I extreme value (Gumbel distribution) error term. The size of a parameter is an estimate of how a change in the variable in question changes the propensity for individuals and organizations to be matched. Our model estimates only parameters that distinguish between matching alternatives (*44*–*46*). An individual's properties may affect the probability of a move, but they do not differentiate between alternative matching opportunities. For this reason, properties of individuals (W_{at}) are entered only as interactions with the variables measuring the organizational properties and network ties in Π_{aijt+1}.

We specify Π_{aijt+1} by including the same organizational-level variables (X) that we used in our dynamic matched sampling procedure, which have been shown to be key drivers for job acquisition processes, together with individual-level interactions (W) and network variables (Z) (see Supplementary Text for details).

The probability that individual a in organization i is matched with organization j at time $t + 1$, can hence be specified as a conditional logit function of Π_{aijt+1}

$$P_{aijt+1} = \frac{\exp(\Pi_{aijt+1} + \ln(S_{jt}))}{\sum_{l=1}^{L} \exp(\Pi_{ailt+1} + \ln(S_{lt}))}$$

where P_{aijt+1} is the probability that individual a and organization j will be matched at time t+1, L is the number of alternative organizations in the choice set, and $\ln(S_{jt})$ is an offset term that accounts for the fact that larger organizations, on average, have proportionally more positions to offer than small organizations (*44*, *47*).

Given this conditional logit model, the simulation is implemented in as follows. We start with individuals and organizations that are exact replicas of all the individuals and organizations that were in our database in 2000.

1) In 2001, new individuals and organizations enter into the virtual labor market, with exactly the same individual and organizational characteristics as the new entrants had in reality.

2) In 2001, individuals and organizations are removed from the virtual labor market if their corresponding real-world entities left the labor market during that year.

3) For all other individuals that were in the empirical labor market in both 2000 and 2001, the following procedure is carried out:

a) For each individual a, we draw a random sample $C_a = (c_1, c_2, \ldots, c_K)$ from all organizations, where c_1 is the organization where the individual currently worked and $c_2 - c_K$ are randomly drawn organizations. We use $K = 1000$; assigning a large random subset of alternatives, as opposed to a full choice set, ensures computational feasibility without sacrificing empirical fit (see fig. S3).

b) From each individual-specific choice set C_a, we sample two alternatives: the stay-alternative (c_1) with 100% probability and one move-alternative based on the predicted probabilities of the conditional logit model.

c) Last, from the stay/move-choice set, we sample as many mobility events as observed in the real data for that year and update the organizational statistics accordingly.

At time $t \geq 2001$, steps 2 to 4 are repeated.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/16/eabf6730/DC1

REFERENCES AND NOTES

1. D. Tomaskovic-Devey, *Gender & Racial Inequality at Work: The Sources and Consequences of Job Segregation* (Cornell Univ. Press, 1993).
2. T. Petersen, L. A. Morgan, Separate and unequal: Occupation-establishment sex segregation and the gender wage gap. *Am. J. Sociol.* **101**, 329–365 (1995).
3. M. A. Belliveau, Blind ambition? The effects of social networks and institutional sex composition on the job search outcomes of elite coeducational and women's college graduates. *Organ. Sci.* **16**, 134–150 (2005).
4. P. N. Cohen, M. L. Huffman, Individuals, jobs, and labor markets: The devaluation of women's work. *Am. Sociol. Rev.* **68**, 443–463 (2003).
5. M. Charles, D. B. Grusky, *Occupational Ghettos: The Worldwide Segregation of Women and Men* (Stanford Univ. Press, Stanford, CA, 2005).
6. Y. M. Ioannides, L. D. Loury, Job information networks, neighborhood effects, and inequality. *J. Econ. Lit.* **42**, 1056–1093 (2004).
7. M. Granovetter, *Getting a Job: A Study of Contacts and Careers* (University of Chicago Press, Chicago and London, [1974] 1995).
8. W. T. Bielby, J. N. Baron, Men and women at work: Sex segregation and statistical discrimination. *Am. J. Sociol.* **91**, 759–799 (1986).
9. J. R. Elliott, Referral Hiring and Ethnically Homogeneous Jobs: How Prevalent Is the Connection and for Whom? *Soc. Sci. Res.* **30**, 401–425 (2001).
10. L. B. Trimble, J. A. Kmec, The role of social networks in getting a job. *Sociol. Compass* **5**, 165–178 (2011).
11. B. F. Reskin, D. B. McBrier, J. A. Kmec, The determinants and consequences of workplace sex and race composition. *Annu. Rev. Sociol.* **25**, 335–361 (1999).
12. J. R. Elliott, M. Sims, Ghettos and barrios: The impact of neighborhood poverty and race on job matching among blacks and Latinos. *Soc. Probl.* **48**, 341–361 (2001).
13. P. V. Marsden, E. H. Gorman, Social Networks, Job Changes, and Recruitment, in *Sourcebook of Labor Markets: Evolving Structures and Processes*, I. Berg, A. L. Kalleberg, Eds. (Plenum Publishers, New York, 2001), pp. 467–502.
14. B. Rubineau, R. M. Fernandez, Tipping points: The gender segregating and desegregating effects of network recruitment. *Organ. Sci.* **26**, 1646–1664 (2015).
15. B. Rubineau, R. M. Fernandez, Missing links: Referrer behavior and job segregation. *Manag. Sci.* **59**, 2470–2489 (2013).
16. P. M. Blau, A macrosociological theory of social structure. *Am. J. Sociol.* **83**, 26–54 (1977).
17. P. M. Blau, T. C. Blum, J. E. Schwartz, Heterogeneity and intermarriage. *Am. Sociol. Rev.* **47**, 45–62 (1982).
18. M. T. Hallinan, S. S. Smith, The effects of classroom racial composition on students' interracial friendliness. *Soc. Psychol. Q.* **48**, 3–16 (1985).
19. K. Joyner, G. Kao, School racial composition and adolescent racial homophily. *Soc. Sci. Q.* **81**, 810–825 (2000).
20. D. R. James, K. E. Taeuber, Measures of segregation. *Sociol. Methodol.* **15**, 1–32 (1985).
21. M. McPherson, L. Smith-Lovin, J. M. Cook, Birds of a feather: Homophily in social networks. *Annu. Rev. Sociol.* **27**, 415–444 (2001).
22. S. Aral, L. Muchnik, A. Sundararajan, Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proc. Natl. Acad. Sci.* **106**, 21544–21549 (2009).
23. A. Grover, J. Leskovec, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016), pp. 855–864.
24. M. J. Salganik, P. S. Dodds, D. J. Watts, Experimental study of inequality and unpredictability in an artificial cultural market. *Science* **311**, 854–856 (2006).
25. E. E. Bruch, R. D. Mare, Neighborhood choice and neighborhood change. *Am. J. Sociol.* **112**, 667–709 (2006).
26. A. Van de Rijt, D. Siegel, M. Macy, Neighborhood chance and neighborhood change: A comment on Bruch and Mare. *Am. J. Sociol.* **114**, 1166–1180 (2009).

Downloaded from https://www.science.org on February 26, 2025

27. K. Stainback, D. Tomaskovic-Devey, *Documenting Desegregation: Racial and Gender Segregation in Private Sector Employment Since the Civil Rights Act* (Russell Sage Foundation, 2012).
28. R. M. Fernandez, I. Fernandez-Mateo, Networks, race, and hiring. *Am. Sociol. Rev.* **71**, 42–71 (2016).
29. S. F. Reardon, School segregation and racial academic achievement gaps. *RSF* **2**, 34–57 (2016).
30. D. S. Massey, A. B. Gross, K. Shibuya, Migration, segregation, and the geographic concentration of poverty. *Am. Sociol. Rev.* **59**, 425–445 (1994).
31. C. R. Shalizi, A. C. Thomas, Homophily and contagion are generically confounded in observational social network studies. *Sociol. Methods Res.* **40**, 211–239 (2011).
32. H. C. White, *Chains of Opportunity: System Models of Mobility in Organizations* (Harvard Univ. Press, 1970).
33. S. A. Stouffer, Intervening opportunities: A theory relating mobility and distance. *Am. Sociol. Rev.* **5**, 845–867 (1940).
34. M. M. Marini, P.-L. Fan, E. Finley, A. M. Beutel, Gender and job values. *Sociol. Educ.* **69**, 49–65 (1996).
35. E. Cech, B. Rubineau, S. Silbey, C. Seron, Professional role confidence and gendered persistence in engineering. *Am. Sociol. Rev.* **76**, 641–666 (2011).
36. S. J. Correll, Gender and the career choice process: The role of biased self-assessments. *Am. J. Sociol.* **106**, 1691–1730 (2001).
37. C. R. Shalizi, E. McFowland III, Estimating causal peer influence in homophilous social networks by inferring latent locations. arXiv:1607.06565 (2016).
38. V. Veitch, Y. Wang, D. Blei, in *Advances in Neural Information Processing Systems* (2019), pp. 13792–13802.
39. S. M. Iacus, G. King, G. Porro, Causal inference without balance checking: Coarsened exact matching. *Polit. Anal.* **20**, 1–24 (2012).
40. G. S. Becker, *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education* (Univ. Chicago Press, 2009).
41. O. Åslund, O. N. Skans, Will I see you at work? Ethnic workplace segregation in Sweden, 1985–2002. *ILR Rev.* **63**, 471–493 (2010).
42. D. Neal, Industry-specific human capital: Evidence from displaced workers. *J. Labor Econ.* **13**, 653–677 (1995).
43. D. E. Ho, K. Imai, G. King, E. A. Stuart, Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit. Anal.* **15**, 199–236 (2007).
44. M. Ben-Akiva, S. R. Lerman, *Discrete Choice Analysis* (MIT press, 1985).
45. D. McFadden, Conditional Logit Analysis of Qualitative Choice Analysis, in *Frontiers in Econometrics*, P. Zarembka, Ed. (Academic Press, 1974).
46. K. E. Train, *Discrete Choice Methods with Simulation* (Cambridge Univ. Press, 2009).
47. E. E. Bruch, R. D. Mare, Methodological issues in the analysis of residential preferences, residential mobility, and neighborhood change. *Sociol. Methodol.* **42**, 103–154 (2012).
48. M. Bygren, Unpacking the causes of segregation across workplaces. *Acta Sociol.* **56**, 3–19 (2013).
49. E. A. Stuart, Matching methods for causal inference: A review and a look forward. *Stat. Sci.* **25**, 1–21 (2010).
50. D. B. Rubin, Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* **66**, 688–701 (1974).
51. C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, 2006).
52. P. R. Rosenbaum, D. B. Rubin, The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**, 41–55 (1983).
53. G. King, R. Nielsen, Why propensity scores should not be used for matching. *Polit. Anal.* **27**, 435–454 (2019).
54. K. Lewis, The limits of racial prejudice. *Proc. Natl. Acad. Sci.* **110**, 18814–18819 (2013).
55. P. C. Austin, Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat. Med.* **28**, 3083–3107 (2009).
56. S. J. Correll, Constraints into preferences: Gender, status, and emerging career aspirations. *Am. Sociol. Rev.* **69**, 93–113 (2016).
57. R. A. Brands, Cognitive social structures in social network research: A review. *J. Organ. Behav.* **34**, S82–S103 (2013).
58. O. Åslund, O. Nordström Skans, How to measure segregation conditional on the distribution of covariates. *J. Popul. Econ.* **22**, 971–981 (2009).
59. G. S. Becker, *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education* (The University of Chicago Press, Chicago, 1962).
60. P. D. Gottlieb, G. Joseph, COLLEGE-TO-WORK migration of technology graduates and HOLDERS of doctorates within the UNITED STATES. *J. Reg. Sci.* **46**, 627–659 (2006).
61. P. S. Davies, M. J. Greenwood, H. Li, A conditional logit approach to U.S. state-to-state migration. *J. Reg. Sci.* **41**, 337–360 (2002).
62. S. J. Davis, J. C. Haltiwanger, S. Schuh, Job creation and destruction. *MIT Press Books* **1**, (1998).

Acknowledgments: We thank A. Daoud, B. Jarvis, M. Keuschnigg, E. Tapia, and C. Steglich for valuable comments. **Funding:** The research leading to these results has received funding from the Swedish Research Council (DNR 445-2013-7681). **Author contributions:** M.A., F.C., and P.H. designed the research, performed the research, analyzed the data, and wrote the paper. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The data we use derive from Swedish administrative and tax records. They can therefore not be shared. However, access can be requested from Statistics Sweden (see www.scb.se/en/services/guidance-for-researchers-and-universities/). Furthermore, code for reproducing our results can be found at https://github.com/martin-arvidsson/TrojanHorseMechanism. Additional information and data related to this paper may be requested from the authors.

Submitted 11 November 2020
Accepted 1 March 2021
Published 16 April 2021
10.1126/sciadv.abf6730

Citation: M. Arvidsson, F. Collet, P. Hedström, The Trojan-horse mechanism: How networks reduce gender segregation. *Sci. Adv.* **7**, eabf6730 (2021).

Downloaded from https://www.science.org on February 26, 2025