Peer Effects in the Workplace: A Network Approach*

Matthew J. Lindquist[†] Jan Sauermann[‡] Yves Zenou[§]

May 15, 2022

Abstract

We study both endogenous and exogenous peer effects in worker productivity using an explicit network approach. We apply this method to data from an in-house call center of a multinational mobile network operator that include detailed information on individual performance. We find that a 10% increase in average co-worker *current* productivity *increases* worker productivity by 5.3%. A 10% increase in average co-worker *permanent* productivity *decreases* worker productivity by 3.2%. Older workers, low tenure workers, and low-permanent productivity workers respond the most to changes in co-worker productivity. These workers free ride in the presence of co-workers from the top quartile of the distribution of permanent productivity. Counterfactual exercises demonstrate how managers could mitigate the problem of free riding by re-shuffling workers into different co-worker networks.

Keywords: peer effects, endogenous peer effects, exogenous peer effects, social networks, worker productivity.

JEL Classification: J24, M50.

*We thank Jordi Blanes i Vidal, Zafer Büyükkeçeci, Edwin Leuven, Arjan Non, and seminar participants at Copenhagen Business School, Stockholm University, EALE, and the Colloquium on Personnel Economics (COPE) for helpful comments. Lindquist gratefully acknowledges funding from the Swedish Research Council (Vetenskapsrådet). Sauermann gratefully acknowledges funding from the Jan Wallander and Tom Hedelius Foundation (Grant number I2011-0345:1). This is a revised version of our earlier CEPR Discussion Paper entitled "Network Effects on Worker Productivity".

[†]SOFI, Stockholm University, Sweden. E-mail: matthew.lindquist@sofi.su.se.

[‡]IFAU; UCLS; Center for Corporate Performance (CCP), Institute for the Study of Labor (IZA), Bonn; Research Centre for Education and the Labour Market (ROA), Maastricht University. E-mail: jan.sauermann@ifau.uu.se.

[§]Monash University, IFN, IZA and CEPR. E-mail: yves.zenou@monash.edu.

1 Introduction

There is a growing empirical literature demonstrating that a worker's productivity can be affected by the productivity and characteristics of her peers. For a large number of tasks across different occupations, studies using laboratory experiments, field experiments, or register data show that co-workers can exert economically significant effects on their peers via channels often not explicitly created by their firms.[1] Understanding the mechanisms behind these effects and the policy consequences of such results may be of great value to firms and other organizations.

When investigating the potential role of peer effects in the work place, a naive estimation strategy of regressing a focal worker's current productivity on her peers' current productivity leads to biased estimates of peers' influence due to the reflection problem (Manski, 1993). To avoid this and other challenges to the identification of causal peer effects, many studies use pre-determined measures of peer ability or permanent productivity along with essentially random exposure to peers. In doing so, these studies measure the effect of co-workers' permanent productivity on a worker's current productivity (see, e.g., Guryan et al. 2009 and Mas and Moretti 2009). In Manski (1993), these types of peer effects are labeled *exogenous* peer effects.

But we may also be interested in measuring *endogenous* peer effects, i.e., the effect of co-workers' current productivity on a worker's current productivity, especially if we believe that current and permanent productivity potentially represent different mechanisms, such as peer-pressure, learning, or free-riding (Kandel and Lazear, 1992; Mas and Moretti, 2009; Bandiera et al., 2013). Ultimately, we may be interested in obtaining measures of both effects within a single context, which (again) has been shown to be a non-trivial task.[2]

[1]Studies which use direct measures of worker productivity include Hamilton et al. (2003), Falk and Ichino (2006), Mas and Moretti (2009), Bandiera et al. (2010), Guryan et al. (2009), De Grip and Sauermann (2012), De Grip et al. (2016), Horrace et al. (2016), Horton and Zeckhauser (2016), Steinbach and Tatsi (2021), Steinbach (2017), and Battiston et al. (2021). For low-skilled jobs, Cornelissen et al. (2017) find corroborative evidence for wages using representative register data. See Herbst and Mas (2015) for a meta-study on peer effects in worker productivity.

[2]See Mas and Moretti (2009, page 121) for a brief discussion.

The aim of this paper is to shed more light on this important topic by studying both endogenous and exogenous peer effects in the workplace using a network-based approach. Specifically, we estimate the effects of both co-worker *current* productivity and co-worker *permanent* productivity on a worker's own *current* productivity. To do this, we use data from an in-house call center of a multinational mobile network operator that include detailed information on the performance of individual workers. Call center agents are organized in teams and team members sit together at work islands. Workers' pay is fixed and does not depend on current own or team performance. We transform punch-clock data, i.e., the exact time when team members enter and leave the workplace, into data on co-worker networks. A weighted link in one of our co-worker networks is defined by the overlap in work hours on a specific day of workers from the same team.

We then use the linear in means model to estimate both endogenous (current productivity) and exogenous (permanent productivity) peer effects in worker productivity. Bramoullé et al. (2009), Calvó-Armengol et al. (2009), De Giorgi et al. (2010), Lee et al. (2010), Liu and Lee (2010), Patacchini et al. (2017), Lee et al. (2021), and others have shown how the architecture of social networks can be used to identify endogenous peer effects seperately from exogenous peer effects.[3] Loosely speaking, endogenous peer effects and exogenous contextual effects are identified if at least two individuals in the same network have different links (Bramoullé et al., 2009).

The network approach does not, however, guarantee identification of causal peer influences on individual behavior. In our context, we face three sources of potential bias arising from (i) correlated shocks, (ii) simultaneity, and (iii) endogenous network formation. We deal with the last issue by using a design that exploits variation not only in the pool of peers but also in the time that they work together; team assignment and work schedules are set in advance by managers and do not respond to movements in current productivity. We deal with the first two issues using an instrumental variables approach that includes fixed

[3]See Blume et al. (2011, 2015), Graham (2015), Boucher and Fortin (2016), Bramoullé et al. (2020), De Paula (2020), and De Paula and Graham (2020) for recent overviews of the literature on the identification of social interactions and for a set of original contributions on the topic.

effects. The key to our Bramoullé et al. (2009) "friends-of-friends" IV strategy is that our co-worker networks are constructed using punch-clock data, which ensures that the exclusion restriction holds; workers on earlier shifts do not work alongside those working on later shifts. We can, therefore, use the fixed characteristics of co-co-workers as instruments for co-worker current productivity.[4]

We find that workers respond to both co-worker current productivity and to co-worker permanent productivity, but in different ways. A 10% increase in average co-worker *current* productivity induces a 5.3% *increase* in own productivity, suggesting that worker productivity is boosted by either knowledge spillovers or through peer pressure and the contemporaneous work pace. At the same time, a 10% increase in average co-worker *permanent* productivity generates *a loss* of 3.2% in own current productivity, which is likely due to workers' free riding behavior. Thus, we find evidence of both productivity enhancing spillovers and free riding within the same context.

Importantly, we find that not all workers respond to changes in co-worker productivity to the same degree. In our setting, older workers, low tenure workers, and low permanent productivity workers respond the most to changes in co-worker productivity. When exploring free riding, we see that these workers free ride in the presence of workers in the top quartile of co-worker permanent productivity. Using our model estimates, we run counterfactual exercises that demonstrate how managers could potentially mitigate the problem of free riding by re-shuffling workers into different co-worker networks.

Our network approach to studying peer effects in worker productivity is similar in spirit to the approaches used by Horrace et al. (2016) and Steinbach and Tatsi (2021). Horrace et al. (2016) estimate a team production function that transforms labor inputs into outputs and allows for productivity spillovers to work through a well defined peer network. In their empirical example, college basketball, they find rather small effects (in comparison to our own). A 10% increase in current co-worker productivity raises own current productivity by 0.5%.

[4] Caeyers and Fafchamps (2020) demonstrate that the network methodology used in this paper is not susceptible to the exclusion bias discussed in Guryan et al. (2009).

The main differences between our work and that of Horrace et al. (2016) is the fact that we study peer effects from both co-worker current productivity and co-worker permanent productivity. Several authors have discussed this distinction but, to the best of our knowledge, it only appears in Steinbach and Tatsi (2021), who have actually included them both in their econometric model of peer effects. In their empirical setting, building airline cargo pallets in a cargo warehouse, they also find evidence of positive spillover effects in current productivity, but negative spillover effects from working with people with high permanent productivity. However, in their setting, the positive spillover effects clearly dominate. They find that a 10% increase in co-worker current productivity increases worker productivity by 2.6%, compared to 5.3% in our setting.

The main difference between our work and that of Steinbach and Tatsi (2021) is that they rely primarily on the spatial distribution of workers in a warehouse in order to identify peer effects, whereas we rely on the temporal distribution of workers. Spatial distance is an important part of our story: the call center workers in our data are assigned to teams that sit together at fixed worked stations; close proximity facilitates peer effects. But in contrast to Steinbach and Tatsi (2021), we use the added dimension of time (overlapping work hours) to construct our co-worker networks and to causally identify peer effects in worker productivity.

Our contributions to the literature on peer effects in the workplace can be summarized as follows: First, we demonstrate the importance of distinguishing between the effects of current co-worker productivity and permanent co-worker productivity. We show that both may be present in a workplace. Second, we show how researchers and practitioners can apply a network methodology to identify the causal effects of these two distinct phenomena using observational data that most large firms have already collected. Third, we illustrate how a simple network model can be used to advise managers on how best to construct teams of co-workers in order to maximize aggregate productivity.

In the next section of this paper, we present our data and illustrate how they can be used to create a set of explicit co-worker networks. Then, in Section 3, we outline our empirical model and discuss our strategy for identifying causal peer effects in worker productivity.

Our results are present in Section 4. Section 5 demonstrates how are network model can be used by managers to think about counterfactual policy exercises that may increase aggregate worker productivity. Section 6 concludes.

2 Data and network definition

2.1 The Call Center Data and Institutional Setting

We use data from an in-house call center of a multinational mobile network operator.[5] The call center is open Monday through Saturday from 7:00 a.m. to 11:00 p.m. The call center provides services for current and prospective customers and is divided into 5 departments, which are segmented by customer group. Our data are from the largest department, which handles calls from private customers with fixed mobile phone contracts. Call center agents working in this department answer customer calls and make notes in their customer database for documentation. We have daily information on workers' performance as well as exact working times starting from September 1, 2008 and ending on December 28, 2009.

All agents are placed in teams, which are led by a team leader. There are 13 teams in our data with approximately 15 workers on each team (over the course of a month). On average, about 8 workers from each team work on any given weekday. The main purpose of grouping agents into teams is that it facilitates monitoring, evaluation, and coaching by the team leader. Teams are not specialized for specific types of calls nor by specific customer groups. There are no team-based incentives.

All teams work on the same floor of the building. The physical workspace is organized into work islands with up to eight agents at an island. Agents who are part of the same team are seated together at a work island.

[5] Although the data used in this study have not been analyzed before, two previous studies use data from the same call centre to estimate alternative types of spillover effects. De Grip and Sauermann (2012) exploit a field experiment to study spillover effects from an on-the-job training program. De Grip et al. (2016) show that newly hired agents learn faster in the presence of co-workers with higher tenure.

To staff the call center with the right amount of agents at any time, the scheduling department makes predictions about the number and types of customer calls *ex ante*. Based on these predictions, they infer the number of agents needed in a week, and for each 30-minute block of a day. This procedure allows for daily and hourly variation in customer demand. Four weeks ahead of their working week, call agents learn about their exact working hours. These working hours also precisely state when agents can take breaks, e.g., for lunch.

Our panel data include daily information on agents' individual performance. Worker performance is continuously, and automatically, measured by the IT system of the call center. Average handling time, i.e., the time an agent needs (on average) to handle a customer call, is the main performance indicator used to evaluate agents. Management's aim is to reduce costs by reducing average handling time $aht_{i,t}$ without a loss in quality. We define a worker's current productivity as $y^c_{i,t} = \frac{100}{aht_{i,t}}$. A decrease in average handling time can be interpreted as an increase in worker performance.

We link these performance data to data on the exact time agents are present at their workplace. This information is gathered from the turnstiles where agents need to log in when entering and log out when leaving the call center. Agents also need to log in and out when they have breaks. Since agents who belong to the same team sit next to each other while working; two agents who are present at the same time are exposed to each other. Thus, we will use team membership and overlap in daily working hours to define links between co-workers and co-worker networks (more on this below). The panel data also include information on agents' gender, age, and tenure.

Descriptive statistics are shown in Appendix Table A1. Our sample consists of 250 workers and 28,418 worker×day observations. Two-thirds of the workers are female. The average age of a worker is approximately 30 years. Most workers work part-time during the afternoon. Worker performance, $y_{i,t}$, varies quite substantially. One standard deviation is equal to 23% of the mean in average worker productivity (see the lower panel of Table A1) and 29% of the mean productivity in our worker×day panel (see the upper panel of Table A1).

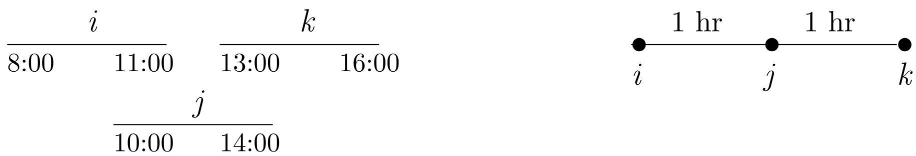
2.2 Defining Co-Worker Networks

There are two levels of co-worker interactions. First, each worker is assigned to a team. Then, each day, individuals work in shifts and interact with different persons within the team that they are allocated to.

Two employees are defined as co-workers if they both come from the same team τ and their work hours on day t overlap. In total, we have 409 workdays of data. We observe 8 or 9 different teams working on a typical day. In total, we observe 3,479 team-by-day networks. To keep the notation simple, we label networks by r, leaving the time and team aspect $r(t, \tau)$ implicit.

We define a weighted adjacency matrix $\mathbf{G}_r$, where each cell $g_{ij,r} \in \{0, h_{i,j}\}$ keeps track of the number of hours $h_{i,j}$ that team members i and j work together on day t. We define the row-normalized matrix $\mathbf{G}^*_r$, such that each cell $g^*_{ij,r} = g_{ij,r}/g_{i,r} \in [0, 1]$, where $g_{i,r} = \sum_j g_{ij,r}$ is the total number of team members' hours worked during agent i's shift on day t. Thus, the weights placed on contemporaneous co-workers sum to one.

Figure 1 illustrates how we use our punch clock data to create co-worker networks. In this example, we have three agents i, j, and k working on day t and who are all members of team τ. Agent i works from 8:00 to 11:00, agent j works from 10:00 to 14:00, and agent k works from 13:00 to 16:00. This means that agent j works one hour together with agent i and one hour with agent k. Agents i and k do not work together on this particular day. This makes worker j the central agent in this particular co-worker network. Agent j acts as a link between agents i and k.



(a) punch clock data (b) network data

Figure 1: Constructing a Co-Worker Network Using Punch Clock Data

The hour-weighted adjacency matrix for this network is:

$$\mathbf{G}_r = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

The row-normalized adjacency matrix is:

$$\mathbf{G}^*_r = \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}.$$

3 Empirical Model

We posit a model of worker productivity that includes both endogenous and exogenous peer effects:

$$\begin{aligned} y_{i,t} = \theta_i + \beta_1 tenure_{i,t} + \beta_2 tenure^2_{i,t} + day_t + shift_s + team_\tau + \\ \phi_1 \mathbf{G}^*_r y_{j,t} + \phi_2 \mathbf{G}^*_r \hat{\mu}_{j,t} + \varepsilon_{i,t,s,\tau}. \end{aligned} \tag{1}$$

Worker i's current productivity on day t, $y_{i,t}$, is determined by several factors. First, each worker i has a fixed effort cost, θ_i. A worker with a low cost of effort will always work harder than a worker with a high cost of effort. We therefore include individual-level fixed effects

θ_i in our empirical model. Second, we observe a strong (and concave) tenure-productivity profile in our data (see Appendix Figure A1). Workers become more productive as they gain experience. In particular, worker productivity rises rapidly during their first six months of employment. We, therefore, add $tenure_{i,t}$ and $tenure^2_{i,t}$ to our empirical model. Third, we add fixed effects for teams, days, and shifts. Some teams may have better or worse managers. Some days and/or shifts may be busier than others, while other days and/or shifts may be hit by positive or negative productivity shocks.

Importantly, we allow workers to be influenced by the average contemporaneous productivity of their co-workers, $\mathbf{G}^*_r y_{j,t}$. In this case, ϕ_1 represents the *endogenous* peer effect. This endogenous effect can, in theory, be either positive or negative. A negative estimate of ϕ_1 implies free riding. A positive effect suggests that there may be strategic complementarities, such as learning, or that there may be peer pressure to perform (and to adhere to the group norm).[6]

A worker's current productivity may also be influenced by the average expected permanent productivity of her co-workers, $\mathbf{G}^*_r \hat{\mu}_{j,t}$. In this case, ϕ_2 represents an *exogenous* peer effect. Workers may be tempted to free ride when working with a group of co-workers who they know (from previous experience) to be high-productivity workers. Alternatively, they could choose to exert more effort if peer pressure increases when exposed to a group of co-workers who are known to be highly productive.

Lastly, we add an individual-level residual term, $\varepsilon_{i,t,s,\tau}$. Standard errors are clustered on individuals, which is more conservative than clustering on networks.

[6]This can be easily derived as a linear best-reply function of a network game with strategic complements. See e.g., Ballester et al. (2006), Patacchini and Zenou (2012), Boucher (2016), Ushchev and Zenou (2020) and, for an overview, Jackson and Zenou (2015).

3.1 Estimating a co-worker's expected permanent productivity

The expected permanent productivity of co-worker j at time t is given by $\mu_{j,t}$. We obtain an estimate of $\mu_{j,t}$ by first estimating:

$$\begin{aligned} y_{j,t} = \theta_j + \alpha_1 tenure_{j,t} + \alpha_2 tenure^2_{j,t} + day_t + shift_s + team_\tau + \\ team_\tau \times day_t + \epsilon_{j,t,s,\tau}. \end{aligned} \tag{2}$$

Equation (2) is based on model (1), but expressed from the perspective of all j co-workers. Estimating Equation (2) produces estimates of the individual fixed effects, $\hat{\theta}_j$, and of the coefficients that describe the average tenure profile in our data, $\hat{\alpha}_1$ and $\hat{\alpha}_2$. The key difference between Equations (1) and (2) is that in Equation (2), we estimate $\hat{\theta}_j$ net of any potential social spillover effects, which is why we include team-by-day (= network) fixed effects.

Worker j's expected permanent productivity, $\hat{\mu}_{j,t}$, is calculated as follows:

$$\hat{\mu}_{j,t} = \hat{\theta}_j + \hat{\alpha}_1 tenure_{j,t} + \hat{\alpha}_2 tenure^2_{j,t}. \tag{3}$$

The main difference between our measure of expected permanent productivity and Mas and Moretti's (2009) measure of permanent productivity is that they examine the effect of permanent co-worker productivity, $\hat{\theta}_j$, only. Here, we also account for the observed tenure profile, which is why we have chosen to call our measure "expected permanent productivity" and not simply "permanent productivity". The underlying idea, however, is the same: worker i knows *ex ante* if worker j is a high or low productivity worker. The distribution of $\hat{\mu}_j$ is shown in Appendix Figure A2.

Since $\hat{\theta}_{j,t}$ included in equation (1) is constructed using estimated parameters provided by Equation (2), we also report bootstrapped standard errors of $\hat{\phi}_2$ in Equation (1) for our baseline results reported in Table 1. The full empirical model is re-estimated in each round of the bootstrap procedure.

3.2 Identification

When estimating peer effects using a linear-in-means model, endogenous and exogenous effects cannot always be separately identified due to the reflection problem (Manski, 1993). When individuals are influenced by the members of their own group, but not by individuals outside their group, there arises a simultaneity in the behavior of individuals within the group that introduces a perfect collinearity between the endogenous peer effect and the exogenous (contextual) peer effects. Using the terminology of social networks, the reflection problem arises when networks are complete. That is, when all agents are connected to (and influenced by) all other agents in the network. However, most networks (such as those studied in this paper) are not complete; everyone is not connected to everyone else. Endogenous peer effects and exogenous contextual effects are identified if at least two individuals in the same network have different links (Bramoullé et al., 2009). This condition is generally satisfied in most real-world networks, including our co-worker networks.

While our network approach does allow us to separately identify endogenous effects and contextual effects, it does not necessarily identify the causal effect of peer influences on individual behavior. In our context, we face three sources of potential bias arising from (i) correlated shocks, (ii) endogenous network formation (non-random sorting), and (iii) simultaneity bias.

Individuals within the same network who share the same environment and face the same set of incentives and/or shocks are likely to behave in a similar manner. We control for these types of correlated effects by adding team, day, and shift (time of day) fixed effects. Recall that networks are defined as team-by-day observations.

Unlike the networks used in most applications (Jackson et al., 2017), our networks are not formed by individuals who self-select into them. Workers do choose to work for the firm and they also state the shifts that they would be willing and able to work. For example, homemakers may want to work in the middle of the day, while students may only be available evenings and weekends. But it is the firm that places these workers into teams and sets the weekly work schedule. This schedule is set one month in advance. Since networks are defined as teammates working on the same day (i.e., we use daily networks), $\mathbf{G}^*_{r(t,\tau)}$ is fixed *ex ante* and does not respond to contemporaneous shocks to either $y_{i,t}$ or $y_{j,t}$.[7]

Lastly, in order to deal with potential simultaneity bias, we adopt Bramoullé et al.'s (2009) "friends-of-friends" instrumental variable strategy. Under the assumptions that (i) $\mathbf{G}^*_r$ is conditionally exogenous, and (ii) the identity matrix $\mathbf{I}$, $\mathbf{G}^*_r$, and $\mathbf{G}^*_r\mathbf{G}^*_r$ are linearly independent, they show how $\mathbf{G}^*_r\mathbf{G}^*_r x_{j,r}$ can be used as valid instruments for the endogenous variable $\mathbf{G}^*_r y_{j,t}$ in Equation (1). Importantly, this IV method also assumes that there is no error in the measurement of network links, which is usually not the case in (for example) friendship networks (Bramoullé et al., 2020). Our use of actual punch clock data implies that our networks are measured without such error.

We apply this instrumental variables strategy to our punch-clock network data together with the fact that peer influences can only flow forward through time and not backward. Recall agents i, j, and k from Figure 1. In this example, we use x_i, the characteristics of individual i working the early shift, as an instrument for y_j, the productivity of individual j working the middle shift, in order to study the influence that y_j may have on y_k, the productivity of individual k working the later shift. Since working during a given shift is not self-reported (as it is for friendships) and can be accurately measured with our punch clock data, and since individual i cannot directly interact with individual k, x_i becomes a valid instrument for y_j and the exclusion restriction holds by construction.

Our IV estimation strategy also requires that there are at least three individuals in each network. There can be no within-network instruments if there are no indirect links. Furthermore, there are no valid instruments for those working the earliest hours during the day, since no one works before them, and since we rely on the forward flow of time to create our instruments. Together, these two restrictions reduce our estimation sample from 28,418

[7] Bayer et al. (2009) propose a test for conditional randomness in group assignment that we use to support this argument. We create a productivity index for each individual by estimating Equation (1) excluding $\mathbf{G}^*_r y_{j,t}$ and $\mathbf{G}^*_r \hat{\mu}_{j,t}$. We then use these coefficients to predict our productivity index, $\hat{y}_{i,t}$. A simple regression shows a strong, positive association between this index and our measures of endogenous and exogenous peer effects. However, after conditioning on our baseline fixed effects, these associations become precisely estimated zeros. This result supports our claim of conditionally random network formation. See Appendix Table A2 for more details.

worker-by-day observations to 25,699. But we still have the same number of workers (250). Appendix Table A3 shows descriptive statistics for our estimation sample, where we see that they are very similar to those of the full sample reported Appendix Table A1.

4 Empirical Results

4.1 Results

Our first set of results are reported in Table 1. The estimates are produced by estimating Equation (1) with average co-worker current productivity and average co-worker expected permanent productivity, respectively, instrumented using the IV estimation strategy shown in the previous section. In Column (1), we estimate the effect of average co-worker current productivity only, $\mathbf{G}_r^* y_{j,t}$, on own current productivity, $y_{i,t}$, while excluding our measure of average co-worker permanent productivity, $\mathbf{G}_r^* \hat{\mu}_{j,t}$. A one standard deviation increase in co-worker current productivity raises own productivity by 4%. This implies that a 10% increase in co-worker current productivity raises own current productivity by 1.4%, which is line with the findings of Falk and Ichino (2006).

In Column (2), we find a *zero effect* when only looking at the impact of average co-worker *expected* permanent productivity on own current productivity, while excluding our measure of average co-worker current productivity.

Estimates from our preferred model that includes both endogenous (current productivity) and exogenous (permanent productivity) peer effects are reported in Column (3) of Table 1. Here, we see that both effects are large and significant. A one standard deviation increase in average co-worker *current productivity* leads to an *increase* in own productivity by 15%. Thus, a 10% increase in co-worker current productivity induces a 5.3% increase in own productivity. A one standard deviation increase in average co-worker *permanent productivity* induces a *reduction* in own current productivity of 9%. This implies that a 10% increase in average co-worker permanent productivity generates a loss of 3.2% in own current productivity.

Table 1: The Effects of Co-Worker Productivity on Own Productivity.

	Baseline Estimates			Placebo Estimates
	(1)	(2)	(3)	(4)
Effect of co-workers' **current productivity** on own productivity				
$\hat{\phi}_1 \mathbf{G}_r^* y_{j,t,s,\tau} - z$	0.012		0.050	0.000
$S.E.$	(0.006)		(0.019)	(0.030)
p-value	0.04		0.01	0.99
effect size %	4%		15%	0%
Effect of co-workers' expected **permanent productivity** on own productivity				
$\hat{\phi}_2 \mathbf{G}_r^* \hat{\mu}_{j,t} - z$		-0.000	-0.032	-0.000
$S.E.$		(0.001)	(0.012)	(0.018)
p-value		0.70	0.01	0.99
Bootstrapped $S.E.$		[0.003]	[0.015]	
p-value		0.90	0.03	
effect size %		-0%	-9%	-0%
N	25699	25699	25699	23299
Kleibergen-Paap F-stat	130		22	4
Exogeneity test	0.05		0.25	0.12
Mean of $y_{i,t,s,\tau}$	0.34	0.34	0.34	0.34

The dependent variable is worker i's current productivity, $y_{i,t,s,\tau}$. Independent variables labeled z have been standardized with men zero and variance 1. Numbers with % signs indicate effect sizes from a one standard deviation increase in the explanatory variable, set in relation to the mean of the dependent variable. $\mathbf{G}_r^* y_{j,t,s,\tau}$ - z is instrumented with $\mathbf{G}_r^{*2} age_j$ - z, $\mathbf{G}_r^{*2} gender_j$ - z, and $\mathbf{G}_r^{*2} tenure_{j,t}$ - z. The row labeled exogeneity test reports the p-value associated with Hansen's J statistic. All regressions include $tenure_{i,t}$ - z, $tenure^2_{i,t}$ - z, and day_t, $shift_s$, and $team_\tau$ fixed effects. Standard errors (in parentheses) for $\hat{\phi}_1$ and $\hat{\phi}_2$ are clustered on individuals, which is more conservative than clustering on networks. Bootstrapped standard errors [in brackets] are provided for $\hat{\phi}_2$. The full empirical model is re-estimated in each round of the bootstrap procedure. The placebo test in Column (4) estimates the same model as in Column (3), but assigns individual workers to random teams each day. The idea being that fake co-workers should have no measurable affect on a worker's current productivity.

Our results show that both endogenous and exogenous peer effects matter. Both produce sizable changes in a worker's current productivity, albeit in opposite directions. Importantly, the exclusion of either one of the two different types of peer effects pushes the other effect towards zero. Note also that our preferred empirical model, Column (3), has a strong first stage and has a high p-value associated with our test of exogenous instruments.

The results in Table 1 are (of course) ceteris paribus. One should keep this assumption in mind when thinking about our result on free riding. Co-worker current productivity and co-worker expected permanent productivity are highly correlated (0.67), since both depend on co-worker tenure.[8] Thus, a 10% increase in co-worker permanent productivity will tend to raise co-worker current productivity by 6.7%. So the net effect of this increase in co-worker permanent productivity will be $-3.2\% + 0.67 \times 5.3\% = +0.35\%$. This does not mean that free riding is not a problem. It is still generating an aggregate loss in potential productivity. But one could easily miss this result if one were to study current and permanent co-worker productivity separately, i.e., if we had only based our conclusions on either Column (1) or Column (2) in Table 1.

The same logic does not apply to our result for co-worker current productivity. It can be manipulated by an employer (through incentives or monitoring for example) without changing co-worker permanent productivity.

Finally, in Column (4), we present estimates from a placebo test. In this test, we estimate the same model as in Column (3), but assign individual workers to random teams each day. The idea here is that a worker's current productivity should not be affected by the average productivity of a set of fake co-workers, which is confirmed by the zero point estimates reported in Column (4).

4.2 Heterogeneous results

In Table 2, we ask the following question: Do worker's respond differently to different levels of average co-worker expected permanent productivity, $\mathbf{G}^*_{t,\tau}\hat{\mu}_{j,t}$? They clearly do. In Column

[8] See equations (2) and (3), Appendix Figure A1, and Appendix Table A.

(3), we see that workers respond most to co-workers with permanent productivity in the top quartile, while they do not respond negatively to those in the first quartile (see Column (2)). It is the response of workers to the most productive co-workers that drives our baseline results (compare the estimates in Column (3) to those in Column (1) of Table 2).

Which worker-specific characteristics are most related to peer effects? In Columns (2) and (3) of Table 3, we see that *low-tenure workers* respond more strongly to both endogenous and exogenous peer effects than high-tenure workers do. In Columns (4) and (5), we see that workers with *low-expected permanent productivity* are affected more by their peers' productivity than are high-permanent productivity workers. For example, a one standard deviation increase in average co-worker current productivity increases a low-productivity worker's current productivity by 14%, while a one standard deviation increase in average co-worker permanent productivity lowers the current productivity of a low-productivity worker by 11%. Lastly, in Columns (6) and (7), we also see that older workers respond more strongly to peer effects than their younger counterparts do.

Table 2: Heterogeneous Effects: Do Workers Respond Differently to Different Levels of Co-Worker Permanent Productivity?

	Baseline (1)	(2)	(3)
Effect of co-workers' **current productivity** on own productivity			
$\hat{\phi}_1 \mathbf{G}^*_{t,\tau} y_{j,t,s,\tau} - z$	0.050	0.017	0.025
S.E.	(0.019)	(0.009)	(0.012)
p-value	0.01	0.05	0.04
Effect of co-workers' expected **permanent productivity** on own productivity			
$\hat{\phi}_2 \mathbf{G}^*_{t,\tau} \hat{\mu}_{j,t} - z$	-0.032		
S.E.	(0.012)		
p-value	0.01		
Quartile 1 (lowest)		0.010	
S.E.		(0.006)	
p-value		0.09	
Quartile 2			-0.011
S.E.			(0.006)
p-value			0.06
Quartile 3			-0.019
S.E.			(0.011)
p-value			0.08
Quartile 4			-0.032
S.E.			(0.016)
p-value			0.05
N	25699	25699	25699
Kleibergen-Paap F-stat	22	67	45
Exogeneity test	0.25	0.05	0.06
Mean of $y_{i,t,s,\tau}$	0.34	0.34	0.34

The dependent variable is worker i's current productivity, $y_{i,t,s,\tau}$. The explanatory variables labeled z have been standardized with mean zero and variance 1. Quartiles 1 through 4 are dummy variables indicating the level of co-worker expected permanent productivity that a worker is exposed to, where quartile 1 is the lowest level of $\mathbf{G}^*_{t,\tau}\hat{\mu}_{j,t}$. $\mathbf{G}^*_r y_{j,t,s,\tau}$ - z is instrumented with $\mathbf{G}^{*2}_r age_j$ - z, $\mathbf{G}^{*2}_r gender_j$ - z, and $\mathbf{G}^{*2}_r tenure_{j,t}$ - z. The row labeled exogeneity test reports the p-value associated with Hansen's J statistic. All regressions include $tenure_{i,t}$ - z, $tenure^2_{i,t}$ - z, and day_t, $shift_s$, and $team_\tau$ fixed effects. Standard errors (in parentheses) for $\hat{\phi}_1$ and $\hat{\phi}_2$ are clustered on individuals, which is more conservative than clustering on networks.

Table 3: Heterogeneous Effects: Who Responds Most to Co-Worker Productivity?

		Tenure		Perm. Prod. $\hat{\mu}_{i,t}$		Age	
	Baseline	High	Low	High	Low	Older	Younger
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Effect of co-workers' **current productivity** on own productivity							
$\hat{\phi}_1 \mathbf{G}^*_{t,\tau} y_{j,t,s,\tau} - z$	0.050	0.035	0.043	0.019	0.060	0.085	0.014
S.E.	(0.019)	(0.036)	(0.026)	(0.038)	(0.036)	(0.033)	(0.018)
p-value	0.01	0.34	0.10	0.63	0.09	0.01	0.45
effect size %	15%	9%	14%	5%	21%	24%	4%
Effect of co-workers' expected **permanent productivity** on own productivity							
$\hat{\phi}_2 \mathbf{G}^*_{t,\tau} \hat{\mu}_{j,t} - z$	-0.032	-0.020	-0.034	-0.008	-0.043	-0.051	-0.015
S.E.	(0.012)	(0.023)	(0.018)	(0.024)	(0.024)	(0.021)	(0.012)
p-value	0.01	0.38	0.06	0.72	0.07	0.01	0.21
effect size %	-9%	- 5%	-11%	-2%	-15%	-15%	-5%
N	25699	12857	12841	12851	12848	12980	12719
Kleibergen-Paap F-stat	22	6	7	6	4	8	15
Exogeneity test	0.25	0.43	n.a.	0.80	0.10	0.28	0.12
Mean of $y_{i,t,s,\tau}$	0.34	0.37	0.31	0.39	0.28	0.35	0.33

The dependent variable is worker i's current productivity, $y_{i,t,s,\tau}$. High tenure is defined as tenure >= median tenure, which is 59 weeks. Low tenure is defined as below median tenure. High expected permanent productivity is defined as $\hat{\mu}_{j,t}$ >= to the median. Low expected permanent productivity is defined as below median $\hat{\mu}_{j,t}$. Older workers are those whose age is >= to the median, which is 28.2 years old. Younger workers are those younger than the median age. The explanatory variables labeled z have been standardized with mean zero and variance 1. Numbers with % signs indicate effect sizes from a one standard deviation increase in the explanatory variable, set in relation to the mean of the dependent variable. $\mathbf{G}^*_r y_{j,t,s,\tau}$ - z is instrumented with $\mathbf{G}^{*2}_r age_j$ - z, $\mathbf{G}^{*2}_r gender_j$ - z, and $\mathbf{G}^{*2}_r tenure_{j,t}$ - z. The row labeled exogeneity test reports the p-value associated with Hansen's J statistic. All regressions include $tenure_{i,t}$ - z, $tenure^2_{i,t}$ - z, and day_t, $shift_s$, and $team_\tau$ fixed effects. Standard errors (in parentheses) for $\hat{\phi}_1$ and $\hat{\phi}_2$ are clustered on individuals, which is more conservative than clustering on networks.

5 Counterfactual Policy Exercises

Our empirical results underscore the importance of peer effects in worker productivity. However, in our setting, they can be both a boon or a bane. Workers increase their work effort when facing a higher work pace, but lower their effort when working alongside highly productive co-workers. Furthermore, we saw that not all workers are equally susceptible to peer effects, and that older, lower tenure, and lower productivity workers respond the most to both current- and permanent co-worker productivity.

Managers could use this fact to incentivize or motivate these workers by inducing a subset of these workers to increase their effort. However, this cannot be done by simply having them work alongside other workers with high-permanent productivity, since this will induce free riding. Instead, the policy must be targeted directly at those workers who respond most to increases in current productivity.

At the same time, managers (in our setting) need to consider ways to reduce the problem of free riding. This could be done, for example, by making co-worker networks more homogeneous. That is, managers could place workers with high and low levels of permanent productivity on different teams and in different co-worker networks. In what follows, we present two counterfactual exercises that illustrate the potential gains that could be achieved by implementing different versions of this policy under different assumptions.

First, we re-estimate our preferred model from Column (3) in Table 1 using variables that are not standardized. Call these non-standardized coefficients $\tilde{\phi}_1$ and $\tilde{\phi}_2$. From the data, we also note that the correlation between average co-worker current productivity, $\mathbf{G}^*_{t,\tau} y_{j,t}$, and average co-worker expected permanent productivity, $\mathbf{G}^*_{t,\tau} \hat{\mu}_{j,t}$, is 0.67. Then, we separate high and low permanent productivity workers from each other. We do this holding average network size constant (= 11 workers), such that there is not a perfect (bimodal) polarization. The goal is to act as a manager and shuffle actual workers across actual networks, while (at the same time) keeping the average network size constant and making sure that we have the necessary personnel needed to fill all days and shifts.

Figure 2 illustrates the consequences of this counterfactual policy. It shows both the new and the old distributions of average co-worker permanent productivity as measured within each separate network. The new distribution is much flatter and the tails are more drawn out. Some networks have quite high average permanent productivity, while others have quite low average permanent productivity, since we have tried to put low permanent productivity workers into different networks than high productivity networks in the hope of mitigating the problem of free riding.

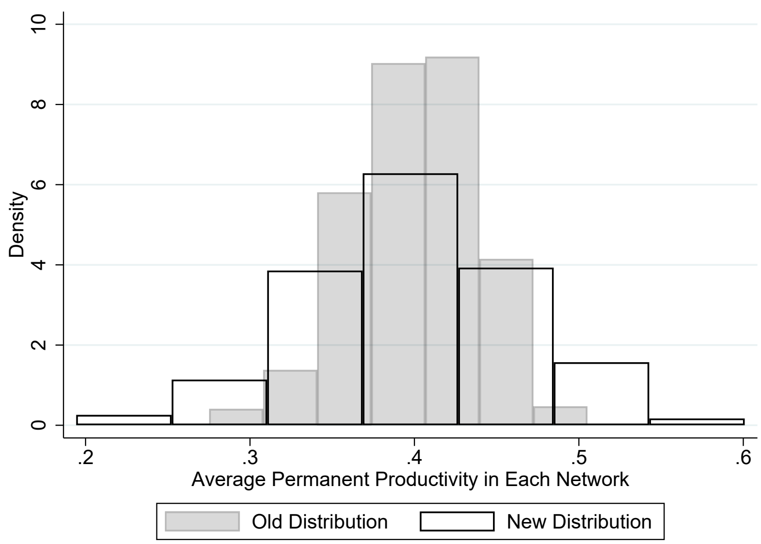


Figure 2: The New and Old Distributions of Average Permanent Productivity in Each Co-Worker Network.

This redistribution of workers changes the average co-worker permanent productivity faced by each worker. This change can be written as $\mathbf{G}^*_{new}\hat{\mu}_{j,t} - \mathbf{G}^*_{old}\hat{\mu}_{j,t}$. This change is positive for some workers and negative for others. The change in average co-worker permanent productivity also induces a change in average co-worker current productivity, since the two measure are correlated (0.67). Now, we can calculate the change in current productivity experienced by each worker i, Δy_i, due to this re-shuffling of the workforce across co-worker networks as follows:

$$\Delta y_i = \tilde{\phi}_1 \times 0.67 \times (\mathbf{G}^*_{new}\hat{\mu}_{j,t} - \mathbf{G}^*_{old}\hat{\mu}_{j,t}) + \tilde{\phi}_2(\mathbf{G}^*_{new}\hat{\mu}_{j,t} - \mathbf{G}^*_{old}\hat{\mu}_{j,t}). \quad (4)$$

The average productivity gain or loss (in %) is then equal to $100 \times \Delta y_i / y_i$ averaged across all i workers. In this counterfactual exercise, the gain from re-shuffling workers in this manner is equal to 0.7%. The source of this gain is a reduction in free riding.

Importantly, one should view this gain as a conservative lower bound. Recall that in Table 3 we saw remarkably different reactions of different types of workers to their co-workers' productivity. The problem of free riding was most apparent among older workers, those with low tenure, and those with low permanent productivity. If we instead allow for heterogeneous parameters (such as those observed in Table 3), we can re-calculate the average productivity gain from this policy as follows:

$$\begin{aligned}\Delta y_{i,K} &= \tilde{\phi}_{1,K} \times 0.67 \times (\mathbf{G}^*_{new}\hat{\mu}_{j,t} - \mathbf{G}^*_{old}\hat{\mu}_{j,t}) + \tilde{\phi}_{2,K}(\mathbf{G}^*_{new}\hat{\mu}_{j,t} - \mathbf{G}^*_{r}\hat{\mu}_{j,t}), \\ &\text{for } K \in \{\text{high permanent productivity, low permanent productivity}\}.\end{aligned} \quad (5)$$

In this version of the counterfactual exercise, we allow for heterogeneous responses from high and low permanent productivity workers. The gain from re-shuffling workers in this manner is now equal to 2.8%, where the source of the gain is (once again) a reduction in free riding.

6 Conclusion

This paper contributes to the literature demonstrating that a worker's productivity can be affected by the productivity and characteristics of her peers. In contrast to most of the existing literature, we estimate the effects of both co-worker current productivity and co-worker permanent productivity on a worker's own current productivity. We find that both matter, but in different ways. A 10% increase in average co-worker *current* productivity induces a 5.3% *increase* in own productivity, while a 10% increase in average co-worker *permanent* productivity generates a *loss* of 3.2% in own current productivity. Thus, we find evidence of both productivity enhancing spillovers and free riding.

Importantly, we find that not all workers respond to changes in co-worker productivity to the same degree. In our setting, it is older workers, low tenure workers, and low permanent productivity workers who respond the most to changes in co-worker productivity. When exploring free riding, we see that these workers free ride in the presence of workers in the top quartile of co-worker permanent productivity. Using our model estimates, we run several counterfactual exercises that demonstrate how managers could potentially mitigate the problem of free riding by re-shuffling workers into different co-worker networks.

References

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006): "Who's who in networks. wanted: the key player," *Econometrica*, 74, 1403–1417.

Bandiera, O., I. Barankay, and I. Rasul (2010): "Social Incentives in the Workplace," *Review of Economic Studies*, 77, 417–459.

——— (2013): "Team Incentives: Evidence From A Firm Level Experiment," *Journal of the European Economic Association*, 11, 1079–1114.

Battiston, D., J. Blanes i Vidal, and T. Kirchmaier (2021): "Face-to-Face Communication in Organisations," *Review of Economic Studies*, 88, 574–609.

Bayer, P., R. Hjalmarsson, and D. Pozen (2009): "Building Criminal Capital behind Bars: Peer Effects in Juvenile Corrections," *The Quarterly Journal of Economics*, 124, 105–147.

Blume, L. E., W. A. Brock, S. N. Durlauf, and Y. M. Ioannides (2011): "Identification of Social Interactions," in *Handbook of Social Economics,*, ed. by J. Benhabib, A. Bisin, and M. O. Jackson, Amsterdam: Elsevier, vol. 1B of *Handbook of Social Economics*, 853–964.

Blume, L. E., W. A. Brock, S. N. Durlauf, and R. Jayaraman (2015): "Linear Social Interactions Models," *Journal of Political Economy*, 123, 444–496.

Boucher, V. (2016): "Conformism and self-selection in social networks," *Journal of Public Economics*, 136, 30–44.

Boucher, V. and B. Fortin (2016): "Some challenges in the empirics of the effects of networks," in *Oxford Handbook on the Economics of Networks*, ed. by B. R. Y. Bramoullé and A. Galeotti, Oxford: Oxford University Press, 277–302.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009): "Identification of peer effects through social networks," *Journal of Econometrics*, 150, 41–55.

Bramoullé, Y., H. Djebbari, and B. Fortin (2020): "Peer Effects in Networks: A Survey," *Annual Review of Economics*, 12, 603–629.

Caeyers, B. and M. Fafchamps (2020): "Exclusion Bias and the Estimation of Peer Effects," CEPR Discussion Paper 14386.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009): "Peer Effects and Social Networks in Education," *The Review of Economic Studies*, 76, 1239–1267.

Cornelissen, T., C. Dustmann, and U. Schönberg (2017): "Peer Effects in the Workplace," *American Economic Review*, 107, 425–56.

De Giorgi, G., M. Pellizzari, and S. Redaelli (2010): "Identification of Social Interactions through Partially Overlapping Peer Groups," *American Economic Journal: Applied Economics*, 2, 241–75.

De Grip, A. and J. Sauermann (2012): "The Effects of Training on Own and Co-Worker Productivity: Evidence from a Field Experiment," *Economic Journal*, 122, 376–399.

De Grip, A., J. Sauermann, and I. Sieben (2016): "Tenure-Performance Profiles and the Role of Peers: Evidence from Personnel Data," *Journal of Economic Behavior & Organization*, 126, 39–54.

De Paula, A. (2020): "Econometric Models of Network Formation," *Annual Review of Economics*, 12, 775–799.

De Paula, A. and B. S. Graham (2020): *The Econometric Analysis of Network Data*, New York: Academic Press.

Falk, A. and A. Ichino (2006): "Clean Evidence on Peer Effects," *Journal of Labor Economics*, 24, 39–58.

Graham, B. S. (2015): "Methods of Identification in Social Networks," *Annual Review of Economics*, 7, 465–485.

Guryan, J., K. Kroft, and M. J. Notowidigdo (2009): "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments," *American Economic Journal: Applied Economics*, 1, 34–68.

Hamilton, B., J. Nickerson, and H. Owan (2003): "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation," *Journal of Political Economy*, 111, 465–497.

Herbst, D. and A. Mas (2015): "Peer effects on worker output in the laboratory generalize to the field," *Science*, 350, 545–549.

Horrace, W. C., X. Liu, and E. Patacchini (2016): "Endogenous network production functions with selectivity," *Journal of Econometrics*, 190, 222–232.

Horton, J. J. and R. J. Zeckhauser (2016): "The Causes of Peer Effects in Production: Evidence from a Series of Field Experiments," Working Paper 22386, National Bureau of Economic Research.

Jackson, M. O., B. Rogers, and Y. Zenou (2017): "The economic consequences of social network structure," *Journal of Economic Literature*, 55.

Jackson, M. O. and Y. Zenou (2015): "Games on Networks," In: P. Young and S. Zamir (Eds.), *Handbook of Game Theory Vol. 4*, Amsterdam: Elsevier Publisher, pp. 91–157.

Kandel, E. and E. P. Lazear (1992): "Peer Pressure and Partnerships," *Journal of Political Economy*, 100, 801–817.

Lee, L.-F., X. Liu, and X. Lin (2010): "Specification and estimation of social interaction models with network structures," *The Econometrics Journal*, 13, 145–176.

Lee, L.-F., X. Liu, E. Patacchini, and Y. Zenou (2021): "Who is the Key Player? A Network Analysis of Juvenile Delinquency," *Journal of Business and Economic Statistics*, 39, 849–857.

Liu, X. and L.-F. Lee (2010): "GMM estimation of social interaction models with centrality," *Journal of Econometrics*, 159, 99–115.

Manski, C. F. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60, 531–542.

Mas, A. and E. Moretti (2009): "Peers at Work," *American Economic Review*, 99, 112–145.

Patacchini, E., E. Rainone, and Y. Zenou (2017): "Heterogeneous peer effects in education," *Journal of Economic Behavior & Organization*, 134, 190–227.

Patacchini, E. and Y. Zenou (2012): "Juvenile delinquency and conformism," *Journal of Law, Economics, and Organization*, 28, 1–31.

Steinbach, D. (2017): "On the Nature of Social Interactions in the Workplace," Ph.D. thesis, Goethe University Frankfurt.

Steinbach, D. and E. Tatsi (2021): "Peer Effects, Free-Riding and Team Diversity," Unpublished manuscript.

Ushchev, P. and Y. Zenou (2020): "Social norms in networks," *Journal of Economic Theory*, 185, 104969.

A Online Appendix

Table A1: Descriptive Statistics for Full Sample

Variable	Mean	Std. Dev.	Min	Max
$y_{i,t}$	.34	.1	.13	1.07
age	31.73	11.02	17.48	60.54
male	.29	.45	0	1
tenure in weeks	133.5	179.46	1	691
weekly hours worked	28	9.14	0	54
morning	.56	.5	0	1
afternoon	.98	.16	0	1
evening	.56	.5	0	1
weekday	.93	.25	0	1
saturday	.07	.25	0	1
N	28418	worker $\times$ day observations		
$y_{i,t}$	.31	.07	.17	.54
age	29.69	10.46	17.48	60.54
male	.32	.47	0	1
tenure in weeks	104.72	170.16	4.04	662.52
weekly hours worked	27.88	6.46	10.48	45.5
morning	.55	.25	0	.97
afternoon	.98	.05	.6	1
evening	.58	.22	.05	.99
weekday	.93	.06	.54	1
saturday	.07	.06	0	.46
N	250	within worker observations		

Table A2: Test of Conditionally Random Group Assignment

Dependent Variable	$\hat{y}_{i,t}$	$\hat{y}_{i,t}$
$\mathbf{G}_r^* y_{j,t}$	1.096	0.000
S.E.	(0.050)	(0.000)
p-value	0.000	0.191
$\mathbf{G}_r^* \hat{\mu}_{j,t}$	1.819	-0.000
S.E.	(0.071)	(0.000)
p-value	0.000	0.402
N	28,418	28,418
Baseline fixed effects	No	Yes

In this table, we present the results of a test for conditionally random group assignment that is similar in sprit to the one proposed by Bayer et al. (2009). We first create a productivity index, $\hat{y}_{i,t}$ for each individual by estimating Equation (1) excluding $\mathbf{G}_r^* y_{j,t}$ and $\mathbf{G}_r^* \hat{\mu}_{j,t}$. We then use these coefficients to predict our productivity index, $\hat{y}_{i,t}$. A simple regression shows a strong, positive association between this index and our measures of endogenous and exogenous peer effects. However, after conditioning on our baseline fixed effects, these associations become precisely estimated zeros. This result supports our claim of conditionally random network formation, since the measures of co-worker productivity are not associated with the predictable parts of an individual worker's productivity once we have controlled for our fixed effects.

Table A3: Summary Statistics for Estimation Sample

Variable	Mean	Std. Dev.	Min	Max
$y_{i,t}$	.34	.1	.13	1.07
age	31.54	10.98	17.48	60.54
gender	.29	.46	0	1
tenure	130.26	178.01	1	691
weekly_hours	28	9.12	0	54
morning	.52	.5	0	1
afternoon	.98	.15	0	1
evening	.6	.49	0	1
weekday	.94	.23	0	1
saturday	.06	.23	0	1
$\mathbf{G}_r^* y_{j,t}$	.34	.06	.14	1.06
$\mathbf{G}_r^* \hat{\mu}_{j,t}$	.4	.04	.23	.6
N	25699	worker x day observations		

Table A4: Correlations

	$y_{i,t}$	$\hat{\mu}_{i,t}$	$age_{i,t}$	$tenure_{i,t}$	$\mathbf{G}_r^* y_{j,t}$	$\mathbf{G}_r^* \hat{\mu}_{j,t}$
$y_{i,t}$	1.00					
$\hat{\mu}_{i,t}$	0.65	1.00				
$age_{i,t}$	0.13	0.14	1.00			
$tenure_{i,t}$	0.24	0.29	0.62	1.00		
$\mathbf{G}_r^* y_{j,t}$	0.42	0.28	0.18	0.26	1.00	
$\mathbf{G}_r^* \hat{\mu}_{j,t}$	0.28	0.41	0.20	0.29	0.67	1.00

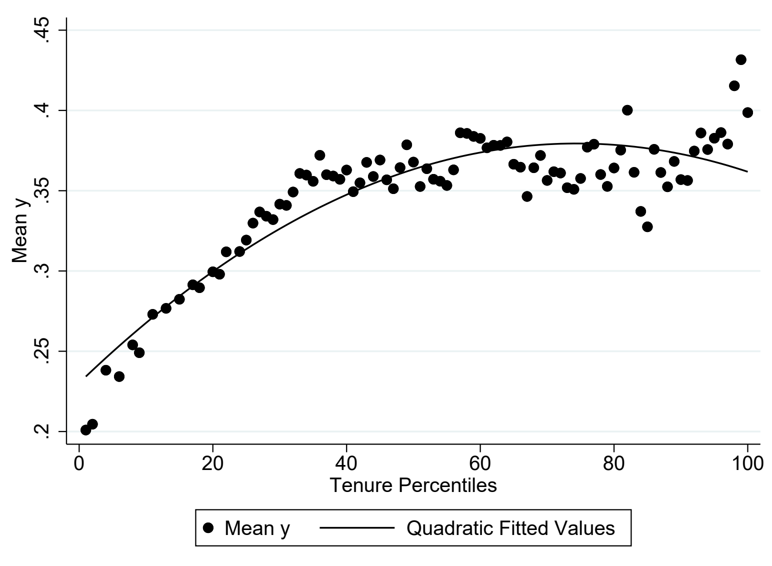


Figure A1: Worker Tenure Productivity Profile.

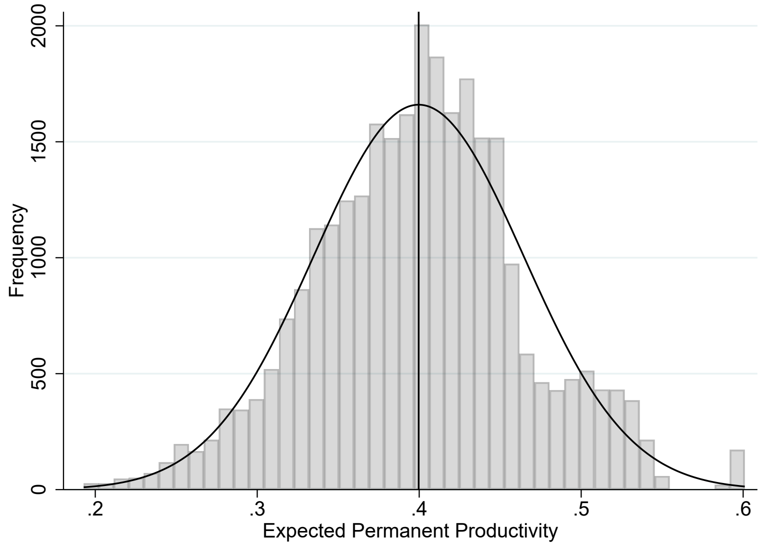


Figure A2: Expected Permanent Productivity, $\hat{\mu}_{j,t}$