DISCUSSION PAPER SERIES

IZA

IZA DP No. 7617

Peer Effects in the Workplace

Thomas Cornelissen
Christian Dustmann
Uta Schönberg

September 2013

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Peer Effects in the Workplace

Thomas Cornelissen
University College London
and CReAM

Christian Dustmann
University College London
and CReAM

Uta Schönberg
University College London,
CReAM, IAB and IZA

Discussion Paper No. 7617
September 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 7617
September 2013

ABSTRACT

Peer Effects in the Workplace*

Existing evidence on peer effects in a work environment stems from either laboratory experiments or from real-word studies referring to a specific firm or specific occupation. Yet, it is unclear to what extent these findings apply to the labor market in general. In this paper, therefore, we investigate peer effects in the workplace for a representative set of workers, firms, and occupations with a focus on peer effects in wages rather than productivity. Our estimation strategy – which links the average permanent productivity of workers' peers to their wages – circumvents the reflection problem and accounts for the endogenous sorting of workers into peer groups and firms. On average, we find only small peer effects in wages. We also find small peer effects in the type of high skilled occupations which more closely resemble those used in studies on knowledge spillover. In the type of low skilled occupations analyzed in existing studies on social pressure, in contrast, we find larger peer effects, about half the size of those identified in similar studies on productivity.

JEL Classification: J24, J31

Keywords: knowledge spillover, social pressure, wage structure

Corresponding author:

Uta Schönberg
Department of Economics
University College London
30 Gordon Street
London WC1H 0AX
United Kingdom
E-mail: u.schoenberg@ucl.ac.uk

* We are grateful to Josh Kinsler for comments and for sharing program code. Uta Schönberg acknowledges funding by the European Research Council (ERC) Starter grant SPILL. Christian Dustmann acknowledges funding by the ERC.

I. Introduction

The communication and social interaction between coworkers that necessarily occur in the workplace facilitate comparison of individual versus coworker productivity. In this context, workers whose productivity falls behind that of coworkers, or falls short of a social norm, may experience personal feelings of guilt or shame. They may then act on these feelings by increasing their own efforts, a mechanism referred to in the economic literature as "peer pressure." Social interaction in the workplace may also lead to "knowledge spillover" in which coworkers learn from each other and build up skills that they otherwise would not have. Both peer pressure and knowledge spillover imply that workers are more productive if their work peers are more productive and that the firm's total productivity exceeds the sum of individual worker productivities. Hence, peer effects, in addition to the transaction cost savings emphasized by Coase (1937), provide one reason for firms' existence. Peer effects may also exacerbate initial productivity differences between workers and increase long-term inequality when high quality workers cluster together in the same peer groups. While knowledge spillover is also an important source of agglomeration economies (e.g., Lucas, 1988; Marshall, 1890), social pressure further implies that workers respond not only to monetary but also to social incentives, which may alleviate the potential free-rider problem inherent whenever workers work together in a team (Kandel and Lazear, 1992).

Yet despite the economic importance of peer effects, empirical evidence on such effects in the *workplace* is as yet restricted to a handful of studies referring to very specific settings, based on either laboratory experiments or on real-world data from a single firm or occupation. For instance, Mas and Moretti's (2009) study of one large supermarket chain provides persuasive evidence that workers' productivity increases when they work alongside more productive coworkers, a finding that they attribute to increased social pressure.

Likewise, a controlled laboratory experiment by Falk and Ichino (2006) reveals that students recruited to stuff letters into envelopes work faster when they share a room than when they sit alone. Other papers focusing on social pressure include Kaur, Kremer, and Mullainathan (2010), who report productivity spillovers among data-entry workers seated next to each other in an Indian company, and Bandiera, Barankay and Rasul (2010), who find that soft-fruit pickers in one large U.K. farm are more productive if at least one of their more able friends is present on the same field, but less productive if they are the most able among their friends.[1]

Turning to studies analyzing peer effects in the workplace due to knowledge spillover, the evidence is mixed. Whereas Azoulay, Graff Zivin, and Wang (2010) and Jackson and Bruegemann (2009) find support for learning from coworkers among medical science researchers and teachers, respectively, Waldinger (2012) finds little evidence for knowledge spillover among scientists in the same department in a university.[2]

While the existing studies provide compelling and clean evidence for the existence (or absence) of peer effects in specific settings, it is unclear to what extent the findings of these studies, which are all based on a specific firm or occupation, apply to the labor market in general. In this paper, we go beyond the existing literature to investigate peer effects in the workplace for a *representative* set of workers, firms, and sectors. Our unique data set, which encompasses all workers and firms in one large local labor market over nearly two decades, allows us to compare the magnitude of peer effects across detailed sectors. It thus provides a rare opportunity to investigate whether the peer effects uncovered in the existing literature are

[1] In related work, Ichino and Maggi (2000) analyze regional shirking differentials in a large Italian bank and find that average peer absenteeism has an effect on individual absenteeism. Furthermore, the controlled field experiment by Babcock et al. (2011) suggests that if agents are aware that their own effort has an effect on the payoff of their peers, this creates incentives. However, this effect is only present for known peers, not for anonymous peers, which suggests that it is mediated by a form of social pressure.

[2] In related work, Waldinger (2010) shows that faculty quality positively affects PhD student outcomes, while Serafinelli (2013) provides evidence that worker mobility from high- to low-wage firms increases the productivity of low-wage firms, which is consistent with knowledge spillover. Other studies (e.g., Guryan, Kroft, and Notowidigdo, 2009; Gould and Winter, 2009) analyze such knowledge spillover between team mates in sports.

confined to the specific firms or sectors studied or whether they carry over to the general labor market, thus shedding light on the external validity of the existing studies. At the same time, our comparison of the *magnitude* of peer effects across sectors provides new evidence on what drives these peer effects, whether social pressure or knowledge spillover.

In addition, unlike the existent studies, our analysis focuses on peer effects in *wages* rather than *productivity*, thereby addressing for the first time whether or not workers are rewarded for a peer-induced productivity increase in the form of higher wages. We first develop a simple theoretical framework in which peer-induced productivity effects arise because of both social pressure and knowledge spillover and translate into peer-related wage effects even when the firm extracts the entire surplus of the match. The rationale behind this finding is that, if the firm wants to ensure that workers remain with the company and exert profit-maximizing effort, it must compensate them for the extra disutility from exerting additional effort because of knowledge spillover or peer pressure.

In the subsequent empirical analysis, we estimate the effect of the long-term or predetermined quality of a worker's current peers—measured by the average wage fixed effect of coworkers in the same workplace (or production site) and occupation—on the current wage, a formulation that directly corresponds to our theoretical model. For brevity, we will from now onwards use the term "firm" to refer to single workplaces or production sites of firms where workers actually work together. We implement this approach using an algorithm developed by Arcidiacono et al. (2012), which allows simultaneous estimation of both individual and peer group fixed effects. Because we link a worker's wage to predetermined characteristics (i.e., the mean wage fixed effect) rather than to peer group wages or effort, we avoid Manski's (1993) reflection problem.

To deal with worker sorting (i.e., that high quality workers may sort into high quality peer groups or firms), we condition on an extensive set of fixed effects. First, by including worker fixed effects in our baseline specification, we account for the potential sorting of high ability workers into high ability peer groups. Further, to account for potential sorting of high ability workers into firms, occupations, or firm-occupation combinations that pay high wages, we include firm-by-occupation fixed effects. To address the possibility that firms may attract better workers and raise wages at the same time, we further include time-variant firm fixed effects (as well as time-variant occupation fixed effects). As argued in Section III.A, this identification strategy is far tighter than most strategies used to estimate peer effects in other settings.

On average, we find only small, albeit precisely estimated, peer effects in wages. This may not be surprising, as many of the occupations in a general workplace setting may not be particularly susceptive to social pressure or knowledge spillover. In fact, the specific occupations and tasks analyzed in the existent studies on peer pressure (i.e., supermarket cashiers, data entry workers, envelope stuffing, fruit picking) are occupations in which there is more opportunity for coworkers to observe each other's output, a prerequisite for peer pressure build-up. Similarly, the specific occupations and tasks analyzed in the studies on knowledge spillover (i.e., scientists, teachers) are high skilled and knowledge intensive, making learning from coworkers particularly important. We therefore restrict our analysis in a second step to occupations similar to those studied in that literature. In line with Waldinger (2012), in occupations for which we expect knowledge spillover to be important (i.e., occupations that are particularly innovative and high skill), we likewise find only small peer effects in wages. On the other hand, in occupations where peer pressure tends to be more important (i.e., where the simple repetitive nature of the tasks makes output more easily observable to coworkers), we find larger peer effects. In these occupations, a 10% increase in peer ability increases wages by 0.6-0.9%, which is about half the size of the effects identified by Falk and Ichino (2006) and Mas and Moretti (2009) for productivity. These findings are remarkably robust to a battery of robustness checks. We provide several types of additional evidence for social pressure being the primary source of these peer effects.

Our results are important for several reasons. First, our finding of only small peer effects in wages *on average* suggests that the larger peer effects established in specific settings in existing studies may not carry over to the labor market in general. Overall, therefore, our results suggest that peer effects do not provide a strong rationale for the existence of firms[3], nor do they contribute much to overall inequality in the economy.

Second, even though our results suggest that the findings of earlier studies cannot be extended to the entire labor market, they also suggest that these earlier findings can be generalized to some extent beyond the single firms or single occupations on which they are based: Our findings highlight larger peer effects in low skilled occupations in which co-workers can, due to the repetitive nature of the tasks performed, easily judge each others' output—which are exactly the type of occupations most often analyzed in earlier studies on peer pressure. Furthermore, our findings add to the existing studies by showing that in such situations, peer effects lead not only to productivity spillover but also to wage spillover, as yet an unexplored topic in the literature.

While being of minor importance for the labor market in general, in the specific sector of low skilled occupations, peer effects do amplify lifetime wage differentials between low and high ability workers. For example, the average peer quality of the 10% most productive workers in these occupations (measured in terms of their fixed worker effect) exceeds the average peer quality of the 10% least productive workers on average by 23%, which, combined with our estimate for peer effects in these occupations, increases the wage differential between these two worker groups by 1.5 to 2%. In comparison, the endogenous sorting of high ability workers into firms or occupations that pay high wages (captured by the

[3] This is generally in line with a recent paper by Bloom et al. (2013) who find that workers who work from home are somewhat more productive than those who come in to work.

firm-year and firm-occupation fixed effects in the regression)—which Card, Heining, and Kline (2013) show to be an important driver of the sharp post-1990 increase in inequality in Germany[4]—exacerbates the wage differential between low and high ability workers in repetitive occupations by about 6%.

The structure of the paper is as follows. The next section outlines a theoretical framework that links peer effects in productivity engendered by social pressure and knowledge spillover to peer effects in wages and clarifies the interpretation of the peer effect identified in the empirical analysis. Sections III and IV then describe our identification strategy and our data, respectively. Section V reports our results, and Section VI summarizes our findings.

II. Theoretical Framework

To motivate our subsequent empirical analysis, we develop a simple principal-agent model of unobserved worker effort in which peer effects in productivity translate into peer effects in wages. In this model, firms choose which wage contract to offer to their employees, providing workers with incentives to exert effort. For any given wage contract, workers are willing to put in more effort if they are exposed to more productive peers either because of social pressure or knowledge spillover, both of which lead to peer effects in productivity. Since within this framework, firms must compensate workers for the cost of effort in order to ensure its exertion, peer effects in productivity will translate into peer effects in wages.

II.A Basic Setup

Production Function and Knowledge Spillover

Consider a firm (the principal) that employs N workers (the agents). In the theoretical analysis, we abstract from the endogenous sorting of workers into firms, which our empirical

[4] Note that Card et al. (2013) investigate only the sorting of high-ability workers into high-wage firms and ignore occupations.

analysis takes into account. We first suppose that worker i produces individual output f_i according to the following production function:

$$f_i = y_i + \varepsilon_i = a_i + e_i(1 + \lambda^K \bar{a}_{\sim i}) + \varepsilon_i,$$

where y_i is the systematic component of worker i's productive capacity, depending on individual ability a_i, individual effort e_i and average peer ability (excluding worker i) $\bar{a}_{\sim i}$, which is included to capture knowledge spillover. It should be noted that in this production function, individual effort and peer ability are complements, meaning that workers benefit from better peers only if they themselves expend effort. In other words, the return to effort is increasing in peer ability, and the greater this increase, the more important the knowledge spillover captured by the parameter λ^K.[5] The component ε_i is a random variable reflecting output variation that is beyond the workers' control and has an expected mean of zero. Firm productivity simply equals the sum of worker outputs. While a worker's ability is exogenously given and observed by all parties, effort is an endogenous choice variable. As is standard in the principal agent literature, we assume that the firm cannot separately observe either worker effort e_i or random productivity shocks ε_i.

Cost of Effort and Social Pressure

Exerting effort is costly to the worker. We assume that in the absence of peer pressure, the cost of effort function is quadratic in effort: $C(e_i) = ke_i^2$. As in Barron and Gjerde (1997), Kandel and Lazear (1992), and Mas and Moretti (2009), we introduce peer pressure by augmenting the individual cost of effort function $C(.)$ with a social "peer pressure" function P(.), which depends on individual effort e_i and average peer output $\bar{f}_{\sim i}$ (excluding

[5] It should be noted that this formulation abstracts from the dynamic implications of knowledge spillover, meaning that the model is best interpreted as one of contemporaneous knowledge spillover through assistance and cooperation between workers on the job. The underlying rationale is that workers with better peers are more productive on the job because they receive more helpful advice from their coworkers than if they were in a low-quality peer group. The existing studies on knowledge spillovers in specific occupations also only look at contemporaneous peers (Azoulay, Graff Zivin, and Wang, 2010; Jackson and Bruegemann, 2009; Waldinger, 2012). Even though knowledge spillovers imply that past peers play a role, one would still expect the current peers to be more important.

worker i). We propose a particularly simple functional form for the peer pressure function: $\left(e_i, \bar{f}_{\sim i}\right) = \lambda^P (m - e_i) \bar{f}_{\sim i}$, where λ^P and m can be thought of as both the "strength" and the "pain" from peer pressure (see below).[6] The total disutility associated with effort thus becomes

$$c_i = C(e_i) + P\left(e_i, \bar{f}_{\sim i}\right) = k e_i^2 + \lambda^P (m - e_i) \bar{f}_{\sim i} .$$

Although the exact expressions derived in this section depend on the specific functional form for the total disutility associated with effort, our general argument does not.

In the peer pressure function, the marginal cost of exerting effort is negative (i.e., $\frac{\partial P(e_i, \bar{f}_{\sim i})}{\partial e_i} = -\lambda^P \bar{f}_{\sim i} < 0$). Thus, workers exert higher effort in the presence of peer pressure than in its absence. The peer pressure function also implies that the marginal cost of worker effort is declining in peer output (i.e., $\frac{\partial^2 P(e_i, \bar{f}_{\sim i})}{\partial e_i \partial \bar{f}_{\sim i}} = -\lambda^P < 0$). In other words, peer quality reduces the marginal cost of effort, and the stronger the peer pressure (captured by λ^P), the larger the reduction. This condition implies that it is less costly to exert an additional unit of effort when the quality of one's peers is high than when it is low. Hence, although peer pressure is often defined by the first condition $\frac{\partial P(e_i, \bar{f}_{\sim i})}{\partial e_i} < 0$ (e.g., Kandel and Lazear, 1992; Mas and Moretti, 2009), it is in fact the second condition $\frac{\partial^2 P(e_i,, \bar{f}_{\sim i})}{\partial e_i \partial \bar{f}_{\sim i}} < 0$ that generates productivity spillover (see also Section II.B). It should further be noted that, for simplicity, we abstract from peer actions like sanctions, monitoring, or punishment, meaning that in our model, peer pressure arises solely through social comparison or "guilt" (Kandel and Lazear, 1992) rather than through sanction, punishment, or "shame."[7]

[6] We assume that $k > \lambda^P$, which not only ensures that the Nash equilibrium is unique (requiring only $2k > \lambda^P$) but also that the firm's maximization problem has an interior solution, see Appendix A.4.

[7] The experimental evidence from Falk and Ichino (2002) indicates that peer pressure can indeed build up from social comparison alone.

It is also worth noting that in our peer pressure function $P(e_i, \bar{f}_{\sim i})$, peer output has a direct effect on worker utility. That is, there is an additional “pain” resulting from higher peer quality, which is governed by the parameter m,[8] on which we impose two bounds in the peer pressure function. First, we require an upper bound for m to ensure that the combined disutility from the direct cost of effort $C(e_i)$ and peer pressure $P(e_i, \bar{f}_{\sim i})$ increases on average in the effort of individual workers in the peer group. Second, like Barron and Gjerde (1997), we assume that m is large enough so that the total cost from peer pressure is increasing in peer quality on average in the peer group. This assumption captures workers’ dislike of working in a high-pressure environment and is a sufficient, albeit not necessary, condition to ensure that peer effects in productivity lead to peer effects in wages. For derivation of the lower and upper bound for m, see Appendices A.1 and A.2.

Wage Contracts and Worker Preferences

Firms choose a wage contract that provides workers with the proper incentives to exert effort. Because the firm cannot disentangle e_i and ε_i, however, it cannot contract a worker’s effort directly but must instead contract output f_i. As it is typical in this literature, we restrict the analysis to linear wage contracts:[9]

$$w_i = \alpha + \beta f_i = \alpha + \beta[a_i + e_i(1 + \lambda^K \bar{a}_{\sim i}) + \varepsilon_i].$$

Contrary to the standard principal agent model, we assume that not only firms but also workers are risk-neutral. This assumption of risk neutrality simplifies our analysis without being a necessary condition for our general argument.

[8] It should be noted that m affects $\frac{\partial P(e_i, \bar{f}_{\sim i})}{\partial \bar{f}_{\sim i}} = \lambda^P(m - e_i)$ but not $\frac{\partial P(e_i, \bar{f}_{\sim i})}{\partial e_i}$ or $\frac{\partial^2 P(e_i, \bar{f}_{\sim i})}{\partial e_i \partial \bar{f}_{\sim i}}$, meaning that the role of m is to mediate the direct effect of peer output on the disutility from peer pressure.

[9] Holmstrom and Milgrom (1987) show that a linear contract is optimal over a range of different environmental specifications.

II.B The Worker's Maximization Problem

Because of risk-neutrality, workers maximize their expected wage minus the combined cost of effort:[10]

$$EU_i = E\left[w_i - C(e_i) - P\left(e_i, \bar{f}_{\sim i}\right)\right] = E[w_i] - C(e_i) - P(e_i, \bar{y}_{\sim i}) \\ = \alpha + \beta[a_i + e_i(1 + \lambda^K \bar{a}_{\sim i})] - ke_i^2 - \lambda^P(m - e_i)\bar{y}_{\sim i}. \quad (1)$$

The maximization problem leads to a linear system of N reaction functions in which each worker in the peer group equates the expected marginal benefit of exerting effort, $\beta(1 + \lambda^K \bar{a}_{\sim i})$, with its expected marginal cost $\frac{\partial C(e_i)}{\partial e_i} + \frac{\partial P(e_i, \bar{y}_{\sim i})}{\partial e_i}$, resulting in the following first order condition (detailed in Appendix A.3):

$$e_i = \frac{\lambda^P}{2k}\bar{e}_{\sim i} + \frac{\beta}{2k} + \frac{\lambda^P + \beta\lambda^K}{2k}\bar{a}_{\sim i} \qquad for \quad i = 1, \dots, N \quad (2)$$

This first order condition not only highlights that equilibrium effort is increasing in peer ability (see last term), either because of peer pressure λ^P or knowledge spillover λ^K, but also that peer pressure ($\lambda^P > 0$) leads to a social multiplier effect whereby the more effort exerted by peers, the more effort exerted by the worker (e_i is increasing in $\bar{e}_{\sim i}$). In sum, both peer pressure and knowledge spillover lead to spillover effects in productivity, a dynamic that incorporates the social multiplier effect arising from peer pressure.

II.C The Firm's Optimization Problem

Firms choose the intercept and slope (or incentive) parameter of the wage contract by maximizing expected profits, $EP = \sum_i (a_i + e_i^*(1 + \lambda^K \bar{a}_{\sim i}) - E[w_i])$, taking into account the workers' optimal effort levels e_i^*, subject to the participation constraint that workers receive a utility that is at least as high as the outside option $v(a_i)$: $EU_i \geq v(a_i)$. As is standard in the

[10] Here we use the fact that $E\left[P\left(e_i, \bar{f}_{\sim i}\right)\right] = P(e_i, \bar{y}_{\sim i})$ because P(.) is linear in $\bar{f}_{\sim i}$, $\bar{f}_{\sim i}$ is linear in $\bar{\varepsilon}_{\sim i}$, and $E[\bar{\varepsilon}_{\sim i}] = 0$. In the following, to simplify notation, we use $P(e_i, \bar{y}_{\sim i})$ in place of $E\left[P\left(e_i, \bar{f}_{\sim i}\right)\right]$.

principal agent literature, we assume that the participation constraint holds with equality, implying that the firm has all the bargaining power and thus pushes each worker to the reservation utility $v(a_i)$. This assumption determines the intercept of the wage contract α as a function of the other model parameters. Solving $EU_i = v(a_i)$ for α, substituting it into the expected wage contract and evaluating it at the optimal effort level yields

$$Ew_i = v(a_i) + C(e_i^*) + P(e_i^*, \bar{y}_{\sim i}), \tag{3}$$

meaning that the firm ultimately rewards the worker for the outside option $v(a_i)$, the cost of effort $C(e_i^*)$, and the disutility from peer pressure $P(e_i^*, \bar{y}_{\sim i})$. We can then derive the firm's first order condition and an expression for the optimal wage contract β^* as detailed in Appendix A.4. In the absence of peer pressure (i.e., $\lambda^P = 0$), we obtain the standard result of an optimal incentive parameter for risk neutral workers that is equal to 1. Interestingly, in the presence of peer pressure, β^* is smaller than 1. Hence, as also noted in Barron and Gjerde (1997), peer pressure constitutes a further reason for the firm to reduce incentives in addition to the well-known trade-off between risk and insurance, which is often emphasized in the principal agent model as important for risk-averse workers.[11]

II.D The Effect of Peer Quality on Wages

How, then, do expected wages depend on average peer ability? In a wage regression that is linear in own ability a_i and peer ability $\bar{a}_{\sim i}$, the coefficient on peer ability approximately identifies the average effect of peer ability on wages, $\frac{1}{N}\sum_i \frac{dEw_i}{d\bar{a}_{\sim i}}$. Differentiating equation (3) and taking averages yields

[11] This outcome results from an externality: the failure of individual workers to internalize in their effort choices the fact that peer pressure causes their peers additional "pain" for which the firm must compensate. The firm mitigates this externality by setting $\beta^* < 1$.

$$\frac{1}{N}\sum_i \frac{dEw_i}{d\bar{a}_{\sim i}} = \underbrace{\frac{1}{N}\sum_i \frac{\partial[C(e_i)+P(e_i,\bar{y}_{\sim i})]}{\partial e_i}\Big|_{\text{optimal}} \underbrace{\frac{de_i^*}{d\bar{a}_{\sim i}}}_{\substack{\text{Productivity}\\ \text{spillover}}}}_{\text{Term 1}} \tag{4}$$

$$+ \underbrace{\frac{1}{N}\sum_i \frac{\partial P(e_i,\bar{y}_{\sim i})}{\partial \bar{y}_{\sim i}}\Big|_{\text{optimal}} \frac{d\bar{y}_{\sim i}}{d\bar{a}_{\sim i}}\Big|_{\text{optimal}}}_{\text{Term 2}}$$

where all terms are evaluated at optimal effort levels and at the optimal β. In Appendix A.5, we show that all terms in (4) are positive. Term 1 consists of a productivity spillover effect due to both knowledge spillover and peer pressure, $\frac{de_i^*}{d\bar{a}_{\sim i}}$, which translates into wages at a rate equal to the marginal cost of effort, $\frac{\partial[C(e_i)+P(e_i,\bar{y}_{\sim i})]}{\partial e_i}\Big|_{\text{optimal}} > 0$. Term 2 (which disappears when there is no peer pressure) captures the fact that higher peer ability is associated with higher peer output ($\frac{d\bar{y}_{\sim i}}{d\bar{a}_{\sim i}}\Big|_{\text{optimal}} > 0$), which causes additional "pain" from peer pressure, ($\frac{\partial P(e_i,\bar{y}_{\sim i})}{\partial \bar{y}_{\sim i}}\Big|_{\text{optimal}} \geq 0$). Our model thus predicts the average effect of peer ability on wages to be unambiguously positive.

III. Empirical Implementation

Next, we describe our estimation strategy for obtaining causal estimates of peer quality on wages that correspond to those in the theoretical analysis. Here, we define a worker's peer group as all workers working in the same (3-digit) occupation and in the same firm in period *t* (see Section IV.B for a detailed discussion of the peer group definition).

III.A Baseline Specification and Identification

We estimate the following baseline wage equation:

$$\ln w_{iojt} = x'_{iojt}\beta + \omega_{ot} + a_i + \gamma\bar{a}_{\sim i,ojt} + \delta_{jt} + \theta_{oj} + v_{iojt} \tag{5}$$

where i indexes workers, o indexes occupations or peer groups, j indexes workplaces or production sites (to which we refer as "firms" for simplicity), and t indexes time periods. Here, $\ln w_{iojt}$ is the individual log real wage, x_{iojt} is a vector of time- variant characteristics with an associated coefficient vector β, ω_{ot} denotes time-variant occupation effects that capture diverging time trends in occupational pay differentials, and a_i is a worker fixed effect. These three latter terms proxy the worker's outside option $v(a_i)$ given in equation (3). The term $\bar{a}_{\sim i,ojt}$ is the average worker fixed effect in the peer group, computed by excluding individual i. The coefficient γ is the parameter of interest and measures the spillover effect in wages ($\frac{1}{N}\sum_i \frac{dEw_i}{d\bar{a}_{\sim i}}$ in equation (4)).

Note that the individual and average worker fixed effects a_i and $\bar{a}_{\sim i,ojt}$ in equation (5) are unobserved and must be estimated. We first discuss the conditions required for a causal interpretation of the peer effect γ assuming that a_i and $\bar{a}_{\sim i,ojt}$ are observed. We then point out the issues that arise from the fact that a_i and $\bar{a}_{\sim i,ojt}$ have to be estimated, in Section III.C.

The individual and average worker fixed effects a_i and $\bar{a}_{\sim i,ojt}$ represent predetermined regressors that characterize a worker's long-term productivity. That is, the peer effect γ in equation (5) captures the reduced-form or total effect of peers' long-term productivity on wages, and embodies not only the direct effect of peer ability on wages (holding peer effort constant), but also the social multiplier effect arising from workers' effort reactions in response to increases in the current effort of their peers. Identification of this effect requires that current peer effort or productivity (or as a proxy thereof, peers' current wages) in equation (5) *not* be controlled for. Thus, in estimating (5), we avoid a reflection problem (Manski, 1993).

Nonetheless, identifying the causal peer effect γ is challenging because of confounding factors such as shared background characteristics. Peer quality may affect a worker's wage simply because high quality workers sort into high quality peer groups or high quality firms, leading to a spurious correlation between peer quality and wages. Our estimation strategy accounts for the endogenous sorting of workers into peer groups or firms by including multiple fixed effects. First, because our baseline specification in equation (5) includes worker fixed effects, it accounts for the potential sorting of high ability workers into high ability peer groups. Second, our inclusion of time-variant firm fixed effects δ_{jt} controls for shocks that are specific to a firm. For example, when bad management decisions result in loss of market share and revenue, wages in that firm may increase at a slower rate than in other firms, motivating the best workers to leave. Therefore, failing to control for time-variant firm fixed effects could induce a spurious correlation between individual wages and peer ability. Third, by controlling for firm-specific occupation effects θ_{oj}, we allow for the possibility that a firm may pay specific occupations relatively well (or badly) compared to the market. For instance, firm A might be known for paying a wage premium to sales personnel but not IT personnel, while firm B is known for the opposite. As a result, firm A may attract particularly productive sales personnel, while firm B may attract particularly productive IT personnel. Hence, once again, failing to control for firm-specific occupation effects could induce a spurious association between individual wages and peer quality.

This identification strategy exploits two main sources of variation in peer quality to estimate the causal effect of peer quality on wages, γ. First, it uses changes in peer quality for workers who remain with their peer group as coworkers join or leave, and that are unexplainable by the overall changes in peer quality occurring in the firm or in the occupation. Second, it exploits changes in peer quality for workers who switch peer groups (after having controlled for the accompanying changes in firm- and occupation-specific fixed effects).

The key identification assumption, provided both the individual and average worker fixed effects a_i and $\bar{a}_{\sim i,ojt}$ are observed, is that—conditional on time-variant control variables, individual permanent ability, occupation-specific time trends, firm-specific time trends and occupation effects—any remaining shocks are uncorrelated with the quality of the peer group: $Cov(\bar{a}_{\sim i,ojt}, v_{iojt} | x_{iojt}^T, a_i, \omega_{ot}, \delta_{jt}, \theta_{oj}) = 0$. It is worth noting that this assumption is considerably weaker than the assumptions typically invoked in the education literature, which seek to identify exogenous spillover effects (e.g., the impact of the share of girls, blacks, immigrants, or grade repeaters on individual performance). For instance, the most common approach in these studies—which measures peer characteristics at the grade level and exploits within-school variation over time (e.g., Gould, Lavy, and Paserman, 2009; Hanushek et al., 2003; Hoxby, 2000; Lavy and Schlosser, 2011; Lavy, Paserman, and Schlosser, 2012)— does not allow for the possibility that the average quality of students (in our case: workers) in the school (in our case: firm) changes over time, or that the effect of the school on students' performance (in our case: wages) may vary over time. An alternative approach in that research measures peer characteristics at the classroom level and exploits within-school grade-year variation (e.g., Ammermueller and Pischke, 2009; Angrist and Lang, 2004; Betts and Zau, 2004; McEwan, 2003; Vigdor and Nechyba, 2007). This requires random assignment of students into classrooms within the school (equivalent to occupations within a firm), thereby ruling out within-school student tracking. Our analysis, in contrast, can account for nonrandom selection into occupations within firms by including firm-specific occupation effects.[12]

[12] Our identification strategy is in some respects also tighter than that of Burke and Sass (2013) who, using a similar approach, measure peer quality in the classroom by the average peer fixed effect. Unlike us, they do not allow for time-variant school fixed effects (in our case, time-variant firm fixed effects) and have no equivalent for our firm-specific occupation effect. On the other hand, they do include teacher fixed effects, thereby allowing for changes in overall school quality over time driven by changes in teacher quality and the differences it makes in performance across classes within the same school, year, and grade.

III.B Within Peer Group Estimator

One remaining problem may be the presence of time-variant peer group-specific wage shocks that are correlated with shocks to peer group quality, which would violate the identification assumption behind our baseline strategy in equation (5). It is unclear a priori whether the existence of such shocks will lead to an upward or downward bias in the estimated peer effect. On the one hand, occupations for which labor demand increases relative to other occupations in the firm may raise wages while simultaneously making it more difficult to find workers of high quality, resulting in a downward bias in the estimated peer effect.[13] On the other hand, a firm may adopt a new technology specific to one occupation only, simultaneously raising wages and worker quality in that occupation (relative to other occupations in the firm) and leading to an upward bias in the estimated peer effect.

One way to deal with this problem is to condition on the full set of time-variant peer group fixed effects p_{ojt}. Although this eliminates the key variation in peer ability highlighted previously to identify the causal peer effect γ, this parameter remains identified—because focal worker i is excluded from the average peer group quality. As a result, the average peer group quality of the same group of workers differs for each worker, and $\bar{a}_{\sim i,ojt}$ varies within peer groups at any given point in time, at least if peer groups are small. Using only within-peer group variation for identification yields the following estimation equation:[14]

$$\ln w_{iojt} = x_{iojt}^T\beta + a_i + \gamma\bar{a}_{\sim i,ojt} + p_{ojt} + \varepsilon_{iojt} \tag{6}$$

This within-peer group estimator, although it effectively deals with unobserved time-variant peer group characteristics, uses limited and specific variation in $\bar{a}_{\sim i,ojt}$: As shown in Appendix B, the spillover effect in equation (6) is identified only if peer groups vary in size.

[13] A similar argument is sometimes made for why labor productivity declines during a boom (see e.g., Lazear et al. 2013).

[14] Because the fixed effects δ_{jt}, ω_{ot} and θ_{oj} do not vary within peer groups at any given point in time, we drop them from this specification.

The advantage of being able to control for time-variant shocks to the peer group is thus countered by the disadvantage that only one particular type of variation is used to identify the effect. The within peer group estimator in equation (6) therefore serves as a robustness check only, rather than as our main specification.

III.C Estimation

Whereas our discussion so far assumes that the individual and average worker fixed effects a_i and $\bar{a}_{\sim i,ojt}$ are observed, they are in fact unobserved and must be estimated. Equations (5) and (6) are then non-linear, producing a nonlinear least squares problem. Because the fixed effects are high dimensional (i.e., we have approximately 600,000 firm years, 200,000 occupation-firm combinations, and 2,100,000 workers), using standard nonlinear least squares routines to solve the problem is infeasible. Rather, we adopt the alternative estimation procedure suggested by Arcidiacono et al. (2012), which is detailed in Appendix C.

According to Arcidiacono et al. (2012), if a_i and $\bar{a}_{\sim i,ojt}$ are unobserved, additional assumptions are required to obtain a consistent estimate of γ (see Theorem 1 in Arcidiacono et al., 2012).[15] Most importantly, the error terms between any two observations (v_{iojt} in our equation (5) baseline specification and ε_{iojt} in our equation (6) within-peer group estimator) must be uncorrelated. In our baseline specification, this assumption rules out any wage shocks common to the peer group, even those uncorrelated with peer group quality. The reason why this additional assumption is needed for consistent estimation when a_i and $\bar{a}_{\sim i,ojt}$ are unobserved is that peer group-specific wage shocks not only affect peer group member

[15] Under these assumptions, Arcidiacono et al. (2012) show that γ can be consistently estimated as the sample size grows in panels with a fixed number of time periods, even though the individual worker fixed effects a_i are generally inconsistent in this situation. Hence, the well-known incidental parameters problem, which often renders fixed effects estimators in models with nonlinear coefficients inconsistent, does not apply to this model.

wages but in panels with short *T* also estimated fixed effects, which could lead to a spurious correlation between individual wages and the *estimated* worker fixed effects in the peer group even when the peer group-specific wage shocks are uncorrelated with the *true* worker fixed effects in the peer group. This bias is mitigated, however, by our panel not being particularly short, with 6.1 observations per worker on average. Moreover, to gauge the magnitude of this bias, we have re-estimated our baseline model in equation (5) after adding a random peer-group level shock to the wage. This has hardly any effect on our estimates, suggesting that a bias due to peer group specific wage shocks is small.[16] Finally, the within-peer group estimator of equation (6) directly deals with this problem by completely eliminating peer-group level wage shocks. The fact that we find similar magnitudes of peer effects from the estimation of the baseline model (5) and of the peer-group fixed effects model (6) confirms that wage shocks correlated at the peer group level do not affect our estimation.

IV. Data

Our data set comes from over three decades of German social security records that cover every man and woman in the system observed on June 30 of each year. It therefore includes virtually the whole employed population except for civil servants, the self-employed, and military personnel.

Our data are particularly suited for the analysis because they include identifiers for single production sites or workplaces (which we refer as "firms" for simplicity) as well as detailed occupational codes that distinguish 331 occupations. Such detail allows us to define peer groups of coworkers in the same firm who are likely to interact. We can also observe all workers in each firm, which allows precise calculation of the average peer group characteristics and ensures that our findings are representative of both the firm and the

[16] For the variance of this added shock, we chose the error term variance from a regression of $\hat{p}_{ojt}$ predicted from (6) on all fixed effects included in (5), which is an empirical estimate of the variance of peer-group specific shocks after these fixed effects are controlled for.

workers. Finally, the longitudinal nature of the data set allows us to follow workers, their coworkers, and their firms over time, as required by our identification strategy, which relies on the estimation of firm and worker fixed effects.

IV.A Sample Selection

We focus on the years 1989-2005, and select all workers aged between 16 and 65 in one large metropolitan labor market, the city of Munich and its surrounding districts. Because most workers who change jobs remain in their local labor market, focusing on one large metropolitan labor market rather than a random sample of workers ensures that our sample captures most worker mobility between firms, which is important for our identification strategy of estimating firm and worker fixed effects. Because the wages of part-time workers and apprentices cannot be meaningfully compared to those of regular full-time workers, we base our estimations on full-time workers not in apprenticeship. Additionally, to ensure that every worker is matched with at least one peer, we drop peer groups (firm-occupation-year combinations) with only one worker.

IV.B Definition of the Peer Group

We define the worker's peer group as all workers employed in the same firm and the same 3-digit occupation, the smallest occupation level available in the social security data. Defining the peer group at the 3-digit (as opposed to the 1- or 2-digit) occupation level not only ensures that workers in the same peer group are likely to interact with each other, a prerequisite for knowledge spillover, but also that workers in the same peer group perform similar tasks and are thus likely to judge each other's output, a prerequisite for peer pressure build-up. Occupations at the 2-digit level, in contrast, often lump together rather different occupations. For instance, the 3-digit occupation "cashiers" belongs to the same 2-digit occupation as "accountants" and "data processing specialists" whose skill level is higher and who perform very different tasks. Defining peer groups at the 3-digit level, on the other hand, increases the variation in peer quality within firms (exploited by our baseline specification) as well as within peer groups (exploited by the within-peer group specification, which vanishes as peer groups become large).

Nevertheless, although this definition seems a natural choice, we recognize that workers may also learn and feel peer pressure from coworkers outside their occupation (i.e., relevant peers omitted from our peer group definition) or may not interact with and/or feel peer pressure from all coworkers inside their occupation (i.e., irrelevant peers included in our peer group definition). In Appendix D, we show that omitting relevant peers generally leads to a downward bias of the true peer effect. Including irrelevant workers, in contrast, causes no bias as long as workers randomly choose with which workers in their occupation to interact. The basic intuition for this surprising result is that the average quality of the observed peers perfectly predicts the average quality of the true peers. Hence, to the extent that workers also learn or feel peer pressure from coworkers outside their 3-digit occupation, our estimates are best interpreted as lower bounds for the true peer effect. In practice, we obtain similar results regardless of whether the peer group is defined at the 2- or 3-digit level (compare column (3) of Table 5 with column (1) Table 4), indicating that our conclusions do not depend on a specific peer group definition.

IV.C Isolating Occupations with High Levels of Peer Pressure and Knowledge Spillover

One important precondition for the build-up of peer pressure is that workers can mutually observe and judge each other's output, an evaluation facilitated when tasks are relatively simple and standardized but more difficult when job duties are diverse and complex. To identify occupations characterized by more standardized tasks, for which we expect peer pressure to be important, we rely on a further data source, the 1991/92 wave of the Qualification and Career Survey (see Gathmann and Schönberg, 2010, for a detailed description). In addition to detailed questions on task usage, respondents are asked how frequently they perform repetitive tasks and tasks that are predefined in detail. From the answers, we generate a combined score on which to rank occupations. We then choose the set of occupations with the highest incidence of repetitive and predefined tasks, which encompasses 5% of the workers in our sample (see column (1) of Appendix Table A1 for a full list of the occupations in this group). This group of most repetitive occupations includes agricultural workers, the subject of Bandiera et al.'s (2010) study, and "cashiers," the focus of Mas and Moretti's (2010) study. The remaining occupations are mostly low skilled manual occupations, such as unskilled laborers, packagers, or metal workers.

For robustness, we also estimate peer effects for the exact same occupations as in the existing studies using real-world data—that is, cashiers (Mas and Moretti, 2009), agricultural helpers (Bandiera et al., 2010), and data entry workers (Kaur et al., 2010)—as well as for a handpicked set of low skilled occupations in which, after initial induction, on-the-job learning is limited. This subgroup, which includes waiters, cashiers, agricultural helpers, vehicle cleaners, and packagers among others, makes up 14% of the total sample (see column (2) of Appendix Table A1 for a full list). Unlike the 5% most repetitive occupations, this group excludes relatively skilled crafts occupations in which learning may be important, such as ceramic workers or pattern makers.

To isolate occupations in which we expect high knowledge spillover, we select the 10% most skilled occupations in terms of workers' educational attainment (average share of university graduates), which includes not only the scientists, academics and teachers used in previous studies (Azoulay et al., 2010; Waldinger, 2012; Jackson and Bruegemann, 2009) but also architects and medical doctors for example. As a robustness check, we also construct a combined index based on two additional items in the Qualification and Career Survey: whether individuals need to learn new tasks and think anew, and whether they need to experiment and try out new ideas, and we pick the 10% of occupations with the highest scores. These again include scientists and academics, but also musicians and IT specialists. We further handpick a group of occupations that appear to be very knowledge intensive, including doctors, lawyers, scientists, teachers, and academics (see columns (3) to (5) of Table A1 for a full list of occupations in these three groups).

It should be noted that when focusing on occupational subgroups, we still estimate the model on the full sample and allow the peer effect to differ for both the respective subgroups and the remaining occupations. Doing so ensures that we use all information available for firms and workers, which makes the estimated firm-year and worker fixed effects—and hence the measure for average peer quality—more reliable.

IV.D Wage Censoring

As is common in social security data, wages in our database are right censored at the social security contribution ceiling. Such censoring, although it affects only 0.7% of the wage observations in the 5% most repetitive occupations, is high in occupations with high expected knowledge spillover. We therefore impute top-coded wages using a procedure similar to that employed by Card et al. (2013) (see Appendix D for details). Whether or not we impute wages, however, our results remain similar even in the high skilled occupations with high censoring. This finding is not surprising given that censoring generally causes the distributions of both worker fixed effects and average peer quality to be compressed in the same way as the dependent variable, meaning that censoring need not lead to a large bias in the estimated peer effect.[17]

[17] In a linear least squares regression with normally distributed regressors, censoring of the dependent variable from above leads to an attenuation of the regression coefficients by a factor equal to the proportion of

IV.E Descriptive Statistics

In Table 1, we compare the 5% most repetitive occupations, in which we expect particularly high peer pressure, and the 10% most skilled occupations, in which we expect high knowledge spillover, against all occupations in our sample. Clearly, the 5% most repetitive occupations are low skilled occupations: nearly half (47%) the workers have no post-secondary education (compared to 17% in the full sample and 4% in the skilled occupations sample) and virtually no worker has graduated from a college or university (compared to 18% in the full sample and 80% in the skilled occupations sample). Moreover, the learning content in the 5% most repetitive occupations is low, while it is high in the 10% most skilled occupations, as implied by responses to whether individuals need to learn new tasks or to experiment with new ideas. The need to cooperate with coworkers, although slightly higher in the skilled sample, is similar in all three samples, as is the median peer group size of 3 or 4 workers per peer group. Not surprisingly, peer group size is heavily skewed, with the mean peer group size exceeding the median peer group size by a factor of about 3-4 in the three samples.

For us to successfully identify peer effects in wages, individual wages must be flexible enough to react to changes in peer quality. Obviously, if firms pay the same wage to workers with the same observable characteristics in the same peer group irrespective of individual productivity, it will be impossible to detect spillover effects in wages even when there are large spillover effects in productivity. According to Figure 1 and the bottom half of Table 1, however, the wages of workers with the same observable characteristics in the same peer

---

uncensored observations (Greene, 1981). Hence censoring the top 15% of observations of the dependent variable attenuates the coefficients by a factor of .85. (This effect of censoring is analogous to the effect of multiplying the dependent variable by .85, which would also attenuate the coefficients by the same factor.) In a model of the form $\ln w_{it} = x_{it}'\beta + a_i + \gamma\bar{a}_{\mathrm{it}} + r_{it}$ (a stylized version of our baseline specification (5)), we would therefore expect the parameters that enter the model linearly, β and a_i (and hence also $\bar{a}_{\mathrm{it}}$), to be attenuated. But given that the variances of $\ln w_{it}$ and $\bar{a}_{\mathrm{it}}$ are both attenuated through censoring in the same way, we would expect the peer effects parameter γ to be unaffected. (This is analogous to multiplying both, the dependent variable and the 'regressor' $\bar{a}_{\mathrm{it}}$ by .85, which would leave the coefficient γ unaffected.)

group are far from uniform: the overall standard deviations of log wages are 0.47 in the full, 0.33 in the repetitive, and 0.37 in the skilled occupations sample, respectively. Importantly, the within-peer group standard deviation of the log wage residuals (obtained from a regression of log wages on quadratics in age and firm tenure and aggregate time trends) is about half the overall standard deviation in the full sample (0.24 vs. 0.47), about two thirds in the 5% most repetitive occupations sample (0.20 vs. 0.33), and about three quarters in the 10% most skilled occupations sample (0.27 vs. 0.37). These figures suggest considerable wage variation among coworkers in the same occupation at the same firm at the same point in time. The last row in Table 1 further reveals that real wages are downwardly flexible: about 9% of workers in the full sample, 4% in the skilled occupations sample, and 13% in the repetitive occupations sample experience a real wage cut from one year to another of at least 5%. Overall, therefore, the results clearly show considerable flexibility in individual wages.

We provide additional information on the structure of our sample in Table 2. Our overall sample consists of 2,115,544 workers, 89,581 firms, and 1,387,216 peer groups. Workers are observed on average for 6.1 time periods and have on average worked for 1.6 firms and in 1.4 different occupations. There are 2.3 peer groups on average per firm and year. In our baseline specification based on equation (5), the standard deviation of the estimated worker fixed effects for the full sample (a_i in equation (5)) is 0.36 or 77% of the overall standard deviation of log wages. The average worker fixed effects in the peer group (excluding the focal worker $\bar{a}_{\sim i,ojt}$ in equation (5)) has a standard deviation of 0.29, which is about 60% of the overall standard deviation of the log wage.

As explained in Section III.A, our baseline specification identifies the causal effect of peers on wages by exploiting two main sources of variation in peer quality: changes to the peer group make-up as workers join and leave the group and moves to new peer groups by the focal worker. In Figure 2, we plot the kernel density estimates of the change in a worker's average peer quality from one year to the next separately for those who remain in the peer group (stayers) and those who leave (movers). Not surprisingly, the standard deviation of the change in average peer quality is more than three times as high for peer group movers than for peer group stayers (0.20 vs. 0.06; see also Table 2). Yet even for workers who remain in their peer group, there is considerable variation in average peer quality from one year to the next, corresponding to roughly 20% of the overall variation in average peer quality. As expected, for peer group stayers, the kernel density has a mass point at zero, corresponding to stayers in peer groups that no worker joins or leaves. In our sample, nearly 90% of peer group stayers work in a peer group with at least some worker turnover. Hence, these workers are likely to experience some change in the average peer quality, even without switching peer groups. At 20%, the average peer group turnover in our sample, computed as 0.5 times the number of workers who join or leave divided by peer group size, is quite large and implies that nearly 20% of workers in the peer group are replaced every year.

V. Results

V.A Baseline results

Table 3 reports the estimates for the impact of average peer quality on wages for the full sample, which covers all workers, firms, and occupations in the one large local labor market. Each column of the table introduces additional control variables to account for shared background characteristics. In column (1), we control only for the worker's own fixed effect (a_i in equation (5)), for quadratics in age and firm tenure (captured by x_{iojt} in equation (5)), and for time-variant occupation fixed effects (ω_{ot} in equation (5)), which proxies for outside options. Although the results suggest that a 10% increase in peer quality increases wages by 4.14%, much of this large "peer effect" is presumably spurious because we have not yet controlled for shared background characteristics. Hence, column (2) also incorporates firm fixed effects, thereby accounting for the possibility that workers employed in better firms, which pay higher wages, are also likely to work with better peers. This inclusion reduces the estimated peer effects by more than half. [18] Allowing the firm fixed effect to vary over time (δ_{jt} in equation (5)) in column (3), reduces the effect only slightly. The results now suggest that a 10% increase in peer quality increases wages by 1.27%.

As discussed in Section III.A, if firms that overpay specific occupations relative to the market also attract better workers into these occupations, then this effect may still reflect shared background characteristics rather than peer causality. To account for this possibility, in column (4), we further control for firm-occupation fixed effects (θ_{oj} in equation (5)), which results in a much smaller estimate: a 10% increase in peer quality now increases the individual wage by only 0.1%. Translated into standard deviations, this outcome implies that a one standard deviation increase in peer ability increases wages by 0.3 percentage points or 0.6 percent of a standard deviation. This effect is about 10–15 times smaller than that identified by Mas and Moretti (2009) for *productivity* among supermarket cashiers in a single firm[19] and about 5–7 times smaller than that reported by Jackson and Bruegemann (2009) for productivity among teachers. Hence, we do not confirm similarly large spillover effects in *wages* for a *representative* set of occupations and firms.

V.B Effects for Occupational Subgroups

Peer Pressure

Even if peer effects are small on *average* for a representative set of occupations, they might still be substantial for specific occupations. Hence, in Panel A of Table 4, we report the results for the 5% most repetitive and predefined occupations (see Appendix Table A1 for a

[18] This specification and the associated estimates are roughly in line with those reported by Lengerman (2002) and Battisti (2012), who also analyze the effects of coworker quality on wages.

[19] In a controlled laboratory study, Falk and Ichino (2006) identify peer effects of similar magnitude as Mas and Moretti (2009).

full list), in which we expect particularly high peer pressure. These occupations also more closely resemble those used in earlier studies on peer pressure. All specifications in the table refer to the baseline specification given by equation (5) and condition on occupation-year, firm-year, and firm-occupation fixed effects, meaning that they correspond to specification (4) in the previous table. For these occupations, we find a substantially larger effect of peer quality on wages than in the full sample: a 10% increase in peer quality raises wages by 0.64% (see column (1)) compared to 0.1% in the full sample (see column (4) of Table 3). Expressed in terms of standard deviation, this implies that a one standard deviation increase in peer quality increases the wage by 0.84%, about half the effect found by Falk and Ichino (2006) and Mas and Moretti (2009) for productivity.

Column (2) of Panel A lists the peer effects for the three occupations used in earlier studies, whose magnitudes are very similar to that for the 5% most repetitive occupations shown in column (1). Column (3) reports the results for the handpicked group of occupations in which we expect output to be easily observable and, following initial induction, limited on-the-job learning. The estimated effect for this occupational group is slightly smaller than that for the 5% most repetitive occupations sample but still about five times as large as the effect for the full occupational sample.

Knowledge Spillover

In Panel B of Table 4, we restrict the analysis to particularly high skilled and innovative occupations with a high scope for learning, in which we expect knowledge spillover to be important. Yet regardless of how we define high skilled occupations (columns (1) to (3)), peer effects in these groups resemble those in the full sample.

Overall, therefore, we detect sizeable peer effects in wages only in occupations characterized by standardized tasks and low learning content, which are exactly the occupations in which we expect peer pressure to matter and which closely resemble the specific occupations investigated in the existent studies on peer pressure.

V.C Robustness Checks

As shown in Table 5, the above conclusions remain robust to a number of alternative specifications. In Panel A we display results for repetitive occupations and peer pressure, and in Panel B for high skilled occupations and knowledge spillover. We report our most important robustness check in column (1), where we implement the within-peer group estimator (see equation (6)) which is not affected by the problems that may be caused by peer-group specific wage shocks possibly correlated with peer group quality (see sections III.B and III.C). In both, repetitive occupations (Panel A) and high skilled occupations (Panel B), the estimated peer effect based on the within-peer specification is very close to the effect derived in the respective baseline specifications (see column (1), Table 4). This similarity of the results provides reassurance that we are indeed picking up a true peer effect rather than a spurious correlation.

The remaining columns in Table 5 report the outcomes of specific changes to the baseline specification from equation (5). In column (2), we report the results when the censored wage observations are not imputed. In column (3), we define the peer group, as well as the 5% most repetitive and 10% most skilled occupations, at the 2-digit rather than the 3-digit occupational level. In column (4), we relax the assumption that the effect of observable characteristics is the same for the repetitive occupations and the high skilled occupations as for all other occupations. For both the repetitive and high skilled occupation samples, all robustness checks yield similar estimates as the baseline estimates reported in Table 4. Hence, having consistently identified sizeable peer effects in wages only in low skilled occupations with repetitive tasks (see Panel A), from here onward, we restrict our analysis to this occupational group.

V.D Peer Pressure or Other Channels?

Knowledge Spillover versus Peer Pressure

Although the low learning content in low skilled occupations seems to suggest social pressure as the most likely cause for the peer effects in the 5% most repetitive occupations, such effects could in principle also be driven by knowledge spillover. Hence, in Panels A and B of Table 6, we provide evidence countering the hypothesis that peer effects in these occupations are driven *only* by knowledge spillover. The first counterclaim posits that in the low skilled occupations we focus on, almost all the on-the-job learning takes place when workers are young or have just joined the peer group. Therefore, in Panel A, we allow the peer effect to differ for older (>35) and younger workers (<=35) (column (1)) and for workers who have been with the peer group for more or less than two years (column (2)). Although we do find that peer effects are larger for younger workers, which is in line with knowledge spillover, we also find positive peer effects for older workers. Moreover, peer effects vary little with tenure in the peer group. Both these findings are difficult to reconcile with peer effects arising from knowledge spillover alone. It should also be noted that although the smaller peer effect for older workers is consistent with knowledge spillover, it is also in line with younger workers responding more strongly to peer pressure or suffering more from the "pain" of peer pressure than older or more experienced workers.

A further important difference between peer effects induced by learning from co-workers and those induced by peer pressure is related to the importance of past peers. If peer effects result from learning, both past and current peers should matter, since the skills learnt from a coworker should be valuable even after the worker or coworker has left the peer group. If peer effects are generated by peer pressure, in contrast, then past peers should be irrelevant conditional on current peers in that workers should feel peer pressure only from these latter. Accordingly, in Panel B of Table 6, we add the average worker fixed effects for the lagged peer group (computed from the estimated worker fixed effects from the baseline model) into our baseline regression. We find that the average quality of lagged peers has virtually no effect on current wages, which further supports the hypothesis that peer pressure is the primary source of peer effects in these occupations.

Team Production versus Peer Pressure

Yet another mechanism that may generate peer effects in wages is a team production technology that combines the input of several perfectly substitutable workers to produce the final good, meaning that the (marginal) productivity of a worker depends on the marginal productivities of the coworkers. To test for the presence of such a mechanism, in Panel C of Table 6, we investigate whether peer effects depend on the degree of cooperation between occupational coworkers. Although we do find larger peer effects in occupations where coworker cooperation matters more (0.081 vs. 0.041), we also identify nonnegligible peer effects in occupations where coworker cooperation is less important, which is difficult to reconcile with peer effects arising from team production alone. It should also be noted that although smaller peer effects in occupations with less coworker cooperation is consistent with team production, it is also in line with peer pressure: workers may be more likely to feel pressure from their peers in occupations that demand more coworker cooperation simply because such cooperation makes it easier to observe coworker output.

Whereas all our previous specifications estimate the effect of *average* peer quality on wages, in Panel D of Table 6, we estimate the effect of the quality of the *top* and *bottom* workers in the peer group on wages. To do so, we split the peer group into three groups: the top 10%, the middle 80%, and the bottom 10% of peers based on the estimated worker fixed effects from our baseline regression.[20] We then regress individual wages on the average worker fixed effect for the three groups, controlling for the same covariates and fixed effects as in our baseline specification and restricting the sample to workers in the middle group. We find that the effect of the average peer quality in the middle group on wages is similar to our baseline effect, while the average productivities of peers in the bottom or top groups have no significant effect on wages. Hence, our baseline peer effects are neither driven entirely by very bad workers nor entirely by very good workers. This observation first rules out a simple chain production model in which team productivity is determined by the productivity of the "weakest link in the chain"; that is, the least productive worker. It further suggests that the peer effects in the 5% most repetitive occupations are not driven solely by the most productive workers in the peer group, even though these latter may increase overall peer group productivity by motivating and guiding their coworkers.[21]

V.E Heterogeneous Peer Effects

Symmetry of Effects

Next, in Table 7, we analyze the possible heterogeneity of peer effects, beginning in Panel A with a test of whether improvements in the average peer group quality have similar effects as deteriorations. To this end, using the peer group stayers, we regress the change in log wages on the change in peer group quality (using the pre-estimated worker fixed effects from our baseline specification) and allow this effect to vary according to whether peer group quality improves or deteriorates (see also Mas and Moretti, 2009, for a similar specification).

[20] Although these shares are quite exact in large peer groups, in small peer groups, the top and bottom do not exactly equal 10%. For example, in a peer group with four workers, one worker falls at the top, one at the bottom, and two in the middle.

[21] In an interesting study in a technology-based services company, Lazear et al. (2012) find that the quality of bosses has significant effects on the productivity of their subordinates. While it might be tempting to interpret the quality of the top 10% of peers in our study as a proxy for boss quality, we prefer not to interpret our findings as informative on boss effects. This is because we cannot ascertain whether more able peers are indeed more likely to become team leaders or supervisors, and bosses do not necessarily have to belong to the same occupation as their subordinates, and hence do not have to be in the same peer group as defined in our data.

Our results show relatively symmetric effects for both improvements and deteriorations. This finding differs somewhat from that of Mas and Moretti (2009), who conclude that positive changes in peer quality matter more than negative changes (see their Table 2, column (4)). Such symmetric effects reinforce our conclusion that peer effects in repetitive occupations are driven mostly by peer pressure, which should increase as peer quality improves and ease as it declines.

Low versus High Ability Workers

In Panel B of Table 7, we explore whether the peer effects in wages differ for low and high ability workers in the peer group (i.e., workers below and above the median in the firm-occupation cell). Like Mas and Moretti (2009), we find that peer effects are almost twice as large for low as for high ability workers. One explanation for this finding is that low ability workers increase their effort more than high ability workers in response to an increase in peer quality (i.e., the peer effect in *productivity* is higher for low than for high ability workers). If this latter does indeed explain peer effect differences between low and high ability workers, then, as Mas and Moretti (2009) stress, firms may want to increase peer group diversity—and maximize productivity—by grouping low ability with high ability workers.

However, our model also suggests an alternative interpretation, namely that low ability workers suffer more from the pain of peer pressure than high ability workers, leading to higher peer effects in *wages* for low than high ability workers, even when peer effects in *productivity* are the same.[22] If such "pain" is the reason for the larger peer effects among low

[22] In our model, low and high ability workers increase their effort by the same amount in response to an increase in peer ability (see equation (A.2)), meaning that the peer effect in productivity is the same for both groups. Given $\frac{dEw_i}{d\bar{a}_{\sim i}}$ in equation (4), this variation across individuals can be explained by $\left.\frac{\partial P(e_i,\bar{y}_{\sim i})}{\partial \bar{y}_{\sim i}}\right|_{\text{optimal}} = \lambda^P(m - e_i^*)$, which is associated with the pain from peer pressure. This term varies inversely with a worker's own optimal effort e_i^*, which in turn varies positively with individual ability (see equation (A.2)), implying that the pain from peer pressure for a given increase in peer ability is higher for low-ability than for high-ability workers.

versus high ability workers, then firms may prefer homogenous peer groups over diverse peer groups because they will save wage costs without lowering productivity.

Males versus Females

We are also interested in determining whether, as some evidence suggests, peer effects in the workplace differ between men and women. For example, in an important paper in social psychology, Cross and Madson (1997) propose the "basic and sweeping" difference that women have primarily interdependent self-schemas that contrast markedly with men's mainly independent ones. As a result, women are more social than men and feel a greater need to belong. If so, we would expect females to respond more strongly to peer pressure or feel greater pain from peer pressure than do males.[23] We do in fact find moderate support for this hypothesis: whereas a 10% increase in peer quality increases wages in the repetitive sector by 0.75% for women, it increases wages for men by only 0.54%, a difference that is significant at the 10% level (Panel C of Table 7).

VI. Conclusions

Although peer effects in the classroom have been extensively studied in the literature (see Sacerdote, 2011, for an overview), empirical evidence on peer effects in the workplace is as yet restricted to a handful of studies based on either laboratory experiments or real-world data from a single firm or occupation. Our study sheds light on the external validity of these existing studies by carrying out a first investigation to date of peer effects in a general workplace setting. Unlike existing studies, our study focuses on peer effects in *wages* rather than in *productivity*.

[23] In our model, the former would be captured by a larger λ^P for women than for men, while the latter would be captured by a larger m for women than for men.

We find only small, albeit precisely estimated peer effects, in wages *on average*. This suggests that the larger peer effects found in existing studies may not carry over to the labor market in general. Yet, our results also reveal sizeable peer effects in low skilled occupations in which co-workers can easily observe each others' output—which are exactly the type of occupations most often analyzed in previous studies on peer pressure. In these types of occupations, therefore, the findings of previous studies extend beyond the specific firms or tasks on which these studies are based. Our findings further show that the productivity spillovers translate into wage spillovers, a dynamic as yet unexplored in the literature.

In the high skilled occupations most often analyzed in studies on knowledge spillover, in contrast, we, like Waldinger (2012), find only small peer effects, similar to those found for the overall labor market. It should be noted, however, that these findings do not necessarily imply that knowledge spillover does not generally matter. First, knowledge spillover in productivity may exceed that in wages. Second, in line with the existing studies on knowledge spillover, our specification assumes that workers learn and benefit from their *current* peers only and ignores the importance of *past* peers. In particular, the knowledge a worker has gained from a coworker may not fully depreciate even when the two no longer work together. Thus, although we expect current peers to be more important than past peers, this approach might underestimate the overall importance of knowledge spillover.

Overall, we conclude that peer effects in the workplace, despite being important in some specific settings, do not importantly affect the wage setting of firms, nor do they contribute significantly to overall inequality in the labor market.

UNIVERSITY COLLEGE LONDON

UNIVERSITY COLLEGE LONDON

UNIVERSITY COLLEGE LONDON AND INSTITUT FÜR ARBEITSMARKT- UND BERUFSFORSCHUNG, NÜRNBERG

References

Abowd, John M., Francis Kramarz, and David N. Margolis. "High wage workers and high wage firms." *Econometrica* 67, no. 2 (1999): 251-333.

Abowd, John M., Robert H. Creecy, and Francis Kramarz. “Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data.” Technical Paper No. TP-2002-06, U.S. Census Bureau, 2002.

Ammermueller, Andreas, and Jörn-Steffen Pischke. "Peer effects in European primary schools: Evidence from the progress in international reading literacy study." *Journal of Labor Economics* 27, no. 3 (2009): 315-348.

Angrist, Joshua D., and Kevin Lang. "Does school integration generate peer effects? Evidence from Boston's Metco Program." *American Economic Review* 94, no. 5 (2004): 1613-1634.

Arcidiacono, Peter, Gigi Foster, Natalie Goodpaster, and Josh Kinsler. "Estimating spillovers using panel data, with an application to the classroom." *Quantitative Economics* 3, no. 3 (2012): 421-470.

Azoulay, Pierre, Joshua Graff Zivin, and Jialan Wang. "Superstar extinction." *Quarterly Journal of Economics* 25 (2010): 549-589.

Babcock, Philip, Kelly Bedard, Gary Charness, John Hartman, Heather Royer. *Letting Down the Team? Evidence of Social Effects of Team Incentives.* Working Paper No. 16687, National Bureau of Economic Research, 2011.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. "Social incentives in the workplace." *Review of Economic Studies* 77, no. 2 (2010): 417-458.

Barron, John M., and Kathy Paulson Gjerde. "Peer pressure in an agency relationship." *Journal of Labor Economics* (1997): 234-254.

Battisti, Michele. "High wage workers and high wage peers." Mimeograph, Simon Fraser University, 2011.

Betts, Julian R., and Andrew Zau. "Peer groups and academic achievement: Panel evidence from administrative data." Working Paper, University of California, San Diego and Public Policy Institute of California, 2004.

Bloom, Nicholas, James Liang, John Roberts, and Zhichun Jenny Ying. *Does Working from Home Work? Evidence from a Chinese Experiment*. Working Paper No. 18871, National Bureau of Economic Research, 2013.

Burke, Mary A., and Tim R. Sass. "Classroom peer effects and student achievement." *Journal of Labor Economics* 31, no. 1 (2013): 51-82.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. "Bootstrap-based improvements for inference with clustered errors." *Review of Economics and Statistics* 90, no. 3 (2008): 414-427.

Card, David, Jörg Heining, and Patrick Kline. “Workplace heterogeneity and the rise of West German wage inequality.” *Quarterly Journal of Economics,* 128 (2013), forthcoming.

Coase, Ronald H. "The nature of the firm." *Economica* 4, no. 16 (1937): 386-405.

Cross, Susan E., and Laura Madson. "Models of the self: Self-construals and gender." *Psychological Bulletin* 122, no. 1 (1997): 5-37.

Falk, Armin, and Andrea Ichino. “Clean evidence on peer effects.” *Journal of Labor Economics*, 24, no. 1 (2006): 39-57.

Gathmann, Christina, and Uta Schönberg. “How general is human capital? A task-based approach.” *Journal of Labor Economics,* 28, no.1 (2010): 1-49.

Gould, Eric D., and Eyal Winter. "Interactions between workers and the technology of production: Evidence from professional baseball." *Review of Economics and Statistics* 91, no. 1 (2009): 188-200.

Gould, Eric D., Victor Lavy, and M. Daniele Paserman. "Does immigration affect the long-term educational outcomes of natives? Quasi-experimental evidence." *Economic Journal* 119, no. 540 (2009): 1243-1269.

Greene, William H. “On the Asymptotic Bias of the Ordinary Least Squares Estimator of the Tobit Model.” *Econometrica* 49, no. 2 (1981): 505-513.

Guryan, Jonathan, Kory Kroft, and Matthew J. Notowidigdo. "Peer effects in the workplace: Evidence from random groupings in professional golf tournaments." *American Economic Journal: Applied Economics* 1, no. 4 (2009): 34-68.

Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivkin. "Does peer ability affect student achievement?" *Journal of Applied Econometrics* 18, no. 5 (2003): 527-544.

Holmstrom, Bengt, and Paul Milgrom. "Aggregation and linearity in the provision of intertemporal incentives." *Econometrica* 55, no. 2 (1987): 303-328.

Hoxby, Caroline. “Peer effects in the classroom: Learning from gender and race variation.” Working Paper No. 7867, National Bureau of Economic Research, 2000.

Ichino, Andrea, and Giovanni Maggi. "Work environment and individual background: Explaining regional shirking differentials in a large Italian firm." *Quarterly Journal of Economics* 115, no. 3 (2000): 1057-1090.

Jackson, Clement Kirabo, and Elias Bruegmann. "Teaching students and teaching each other: The importance of peer learning for teachers." *American Economic Journal: Applied Economics* 1, no. 4 (2009): 85-108.

Kandel, Eugene, and Edward P. Lazear. "Peer pressure and partnerships." *Journal of Political Economy* (1992): 801-817.

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan. "Self-control and the development of work arrangements." *American Economic Review, Papers and Proceedings* 100, no. 2 (2010): 624-628.

Lavy, Victor, and Analia Schlosser. "Mechanisms and impacts of gender peer effects at school." *American Economic Journal: Applied Economics* 3, no. 2 (2011): 1-33.

Lavy, Victor, M. Daniele Paserman, and Analia Schlosser. "Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom." *Economic Journal* 122, no. 559 (2012): 208-237.

Lazear, Edward P., Kathryn L. Shaw, and Christopher T. Stanton. “The Value of Bosses.” Working Paper No. 18317, National Bureau of Economic Research, 2012.

Lazear, Edward P., Kathryn L. Shaw, and Christopher T. Stanton. “Making Do With Less: Working Harder During Recessions.” Mimeo, Stanford University, Graduate School of Business, 2013.

Lengermann, Paul. "Is it who you are, where you work, or with whom you work? Reassessing the relationship between skill segregation and wage inequality." LEHD Technical Paper No. 10, U.S. Census Bureau, Center for Economic Studies, 2002.

Lucas, Robert E., Jr. "On the mechanics of economic development." *Journal of Monetary Economics* 22, no. 1 (1988): 3-42.

Manski, Charles F. "Identification of endogenous social effects: The reflection problem." *Review of Economic Studies* 60, no. 3 (1993): 531-542.

Marshall, Alfred. *Principles of economics*. New York: Macmillan. 1890.

Mas, Alexandre, and Enrico Moretti. “Peers at work.” *American Economic Review* 99, no. 1 (2009): 112-145.

McEwan, Patrick J. "Peer effects on student achievement: Evidence from Chile." *Economics of Education Review* 22, no. 2 (2003): 131-141.

Sacerdote, Bruce. "Peer effects in education: How might they work, how big are they, and how much do we know thus far?" *Handbook of the Economics of Education* 3 (2011): 249-277.

Serafinelli, Michel. “Good firms, worker flows, and productivity.” Mimeo, University of California, Berkeley, 2013.

Vigdor, Jacob, and Thomas Nechyba. Peer effects in North Carolina public schools. In: Woessmann, L., Peterson, P.E. (Eds.), “Schools and the Equal Opportunity Problem” (pp. 73-101). Cambridge, MA: MIT Press, 2007.

Waldinger, Fabian. “Quality Matters: The Expulsion of Professors and the Consequences for PhD Students Outcomes in Nazi Germany.” *Journal of Political Economy,* 118, no. 4 (2010): 787-831, 2010.

Waldinger, Fabian. "Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany." *Review of Economic Studies* 79, no. 2 (2012): 838-861.

APPENDIX A: **Model**

A.1 Derivation of the upper bound of m

For the condition to hold that the combined disutility from the direct cost of effort $C(e_i)$ and peer pressure $P(e_i, \bar{y}_{\sim i})$ in the peer group is on average increasing in individual effort, we require that $\frac{1}{N}\sum_i \frac{\partial[C(e_i)+P(e_i,\bar{y}_{\sim i})]}{\partial e_i} = \frac{1}{N}\sum_i [2ke_i^* - \lambda^P \bar{y}_{\sim i}] > 0$ or equivalently, $(2k - \lambda^P)\bar{e}^* - \lambda^P \bar{a} > 0$. Substituting $\bar{e}^* = \frac{1}{N}\sum_i e_i^* = \frac{\beta^* + \beta^* \lambda^K \bar{a} + \lambda^P \bar{a}}{2k - \lambda^P}$, obtained from the optimal effort levels e_i^* derived in Appendix A.3, gives

$$\beta^* + \beta^* \lambda^K \bar{a} + \lambda^P \bar{a} - \lambda^P \bar{a} > 0$$
$$\beta^* + \beta^* \lambda^K \bar{a} > 0$$
$$\beta^* > 0,$$

implying that only values of *m* that lead to a positive β^* can satisfy this condition. Using β^* derived in equation A.4 in Appendix A.4, the upper bound for *m* is implicitly defined by

$$\frac{\sum_i \frac{\partial e_i^*}{\partial \beta}(1 + \lambda^K \bar{a}_{\sim i}) - \sum_i \lambda^P (m^{upper} - e_i^*) \frac{\partial \bar{e}_{\sim i}^*}{\partial \beta}}{\sum_i \frac{\partial e_i^*}{\partial \beta}(1 + \lambda^K \bar{a}_{\sim i})} = 0.$$

A.2 Derivation of the lower bound of m

Inspection of β^* in equation (A.4) reveals that $\beta^* \leq 1$if $\frac{1}{N}\sum_i \frac{\partial P(e_i,\bar{y}_{\sim i})}{\partial \bar{y}_{\sim i}}\Big|_{\text{optimal}} \frac{\partial \bar{e}_{\sim i}^*}{\partial \beta} \geq 0$, $\Leftrightarrow \frac{1}{N}\sum_i \lambda^P (m - e_i^*) \frac{\partial \bar{e}_{\sim i}^*}{\partial \beta} \geq 0$. In words, the derivative of the cost of peer pressure should be non-decreasing in peer quality on average, where the average is weighted by $\frac{\partial \bar{e}_{\sim i}^*}{\partial \beta}$. If this condition would not hold, then the firm could lower its wage cost by increasing β^* higher than one, because workers on average like the additional peer pressure created by their peers’ higher effort, and would be willing to forgo wages to enjoy this peer pressure. Our assumption rules this case out. The lower bound for m is thus implicitly defined by

$$\frac{1}{N}\sum_i \lambda^P (m^{lower} - e_i^*) \frac{\partial \bar{e}_{\sim i}^*}{\partial \beta} = 0.$$

A.3 Solution to the system of reaction functions

The workers’ first order conditions lead to a linear system of *N* reaction functions:

$\beta(1+\lambda^K\bar{a}_{\sim i}) - \frac{\partial[C(e_i)+P(e_i,\bar{y}_{\sim i})]}{\partial e_i} = 0 \quad for\ \ i = 1,\dots,N$, or

$\beta(1+\lambda^K\bar{a}_{\sim i}) - (2ke_i - \lambda^P\bar{e}_{\sim i} - \lambda^P\bar{a}_{\sim i}) = 0 \quad for\ \ i = 1,\dots,N$, or

$$e_i = \frac{\lambda^P}{2k}\bar{e}_{\sim i} + \frac{\beta}{2k} + \frac{\lambda^P + \beta\lambda^K}{2k}\bar{a}_{\sim i} \qquad for\ \ i = 1,\dots,N$$

Since $k > \lambda^P$, there exists a unique solution to this reaction function system. It should also be noted that $\bar{e}_{\sim i} = \mathrm{N}\bar{\mathrm{e}} - e_i$, meaning it can be rewritten as

$$e_i = \frac{\lambda^P}{2k}\frac{1}{N-1}[\mathrm{N}\bar{\mathrm{e}} - e_i] + \frac{\beta}{2k} + \frac{\lambda^P + \beta\lambda^K}{2k}\bar{a}_{\sim i}$$

Solving for e_i then gives

$$e_i\left(1 + \frac{\lambda^P}{2k}\frac{1}{N-1}\right) = \frac{\lambda^P}{2k}\frac{\mathrm{N}}{N-1}\bar{\mathrm{e}} + \frac{\beta}{2k} + \frac{\lambda^P + \beta\lambda^K}{2k}\bar{a}_{\sim i}$$

$$e_i\left(\frac{2k(N-1) + \lambda^P}{2k(N-1)}\right) = \frac{\lambda^P}{2k}\frac{\mathrm{N}}{N-1}\bar{\mathrm{e}} + \frac{\beta}{2k} + \frac{\lambda^P + \beta\lambda^K}{2k}\bar{a}_{\sim i}$$

$$e_i = \frac{\lambda^P\mathrm{N}}{2k(N-1) + \lambda^P}\bar{\mathrm{e}} + \frac{(N-1)\beta}{2k(N-1) + \lambda^P} + \frac{(\lambda^P + \beta\lambda^K)(N-1)}{2k(N-1) + \lambda^P}\bar{a}_{\sim i} \qquad \text{(A.1)}$$

Taking averages on both sides of this equation yields

$$\bar{\mathrm{e}} = \frac{\lambda^P\mathrm{N}}{2k(N-1) + \lambda^P}\bar{\mathrm{e}} + \frac{(N-1)\beta}{2k(N-1) + \lambda^P} + \frac{(\lambda^P + \beta\lambda^K)(N-1)}{2k(N-1) + \lambda^P}\bar{a}$$

after which solving for $\bar{\mathrm{e}}$ gives

$$\bar{\mathrm{e}}\left(\frac{(2k-\lambda^P)(N-1)}{2k(N-1) + \lambda^P}\right) = \frac{(N-1)\beta}{2k(N-1) + \lambda^P} + \frac{(\lambda^P + \beta\lambda^K)(N-1)}{2k(N-1) + \lambda^P}\bar{a}$$

$$\bar{\mathrm{e}} = \frac{\beta}{(2k-\lambda^P)} + \frac{(\lambda^P + \beta\lambda^K)}{(2k-\lambda^P)}\bar{a}$$

$$\bar{\mathrm{e}} = \frac{\beta}{(2k-\lambda^P)} + \frac{(\beta\lambda^K + \lambda^P)(N-1)}{(2k-\lambda^P)N}\bar{a}_{\sim i} + \frac{(\beta\lambda^K + \lambda^P)}{(2k-\lambda^P)N}a_i$$

Substituting this expression into A.1 yields

$$e_i^* = \frac{\lambda^P \mathrm{N}}{2k(N-1)+\lambda^P}\left[\frac{\beta}{(2k-\lambda^P)} + \frac{(\beta\lambda^K+\lambda^P)(N-1)}{(2k-\lambda^P)N}\bar{a}_{\sim i} + \frac{(\beta\lambda^K+\lambda^P)}{(2k-\lambda^P)N}a_i\right] + \frac{(N-1)\beta}{2k(N-1)+\lambda^P} + \frac{(\lambda^P+\beta\lambda^K)(N-1)}{2k(N-1)+\lambda^P}\bar{a}_{\sim i}$$

$$= \frac{\beta[2k(N-1)+\lambda^P]}{[2k(N-1)+\lambda^P](2k-\lambda^P)} + \frac{\lambda^P(\beta\lambda^K+\lambda^P)(N-1)}{[2k(N-1)+\lambda^P](2k-\lambda^P)}\bar{a}_{\sim i} + \frac{\lambda^P(\beta\lambda^K+\lambda^P)}{[2k(N-1)+\lambda^P](2k-\lambda^P)}a_i + \frac{(\lambda^P+\beta\lambda^K)(N-1)}{2k(N-1)+\lambda^P}\bar{a}_{\sim i},$$

or

$$e_i^* = \frac{\beta[2k(N-1)+\lambda^P] + \lambda^P(\beta\lambda^K+\lambda^P)a_i + 2k(\beta\lambda^K+\lambda^P)(N-1)\bar{a}_{\sim i}}{[2(N-1)k+\lambda^P](2k-\lambda^P)} \quad \text{(A.2)}$$

A.4 The Firm's Maximization Problem

Substituting equation (3) into the profit function produces the following optimization problem for the firm's choice of β:

$$\max_\beta EP = \sum_i [a_i + e_i^*(1+\lambda^K\bar{a}_{\sim i})\bar{a}_{\sim i} - v(a_i) - C(e_i^*) - P(e_i^*, \bar{a}_{\sim i} + \bar{e}_{\sim i}^*)]$$

with first order condition

$$\sum_i \frac{\partial e_i^*}{\partial\beta}(1+\lambda^K\bar{a}_{\sim i}) - \sum_i \left(\frac{\partial C_i}{\partial e_i} + \frac{\partial P_i}{\partial e_i}\right)\frac{\partial e_i^*}{\partial\beta} - \sum_i \frac{\partial P_i}{\partial \bar{e}_{\sim i}^*}\frac{\partial \bar{e}_{\sim i}^*}{\partial\beta} = 0. \quad \text{(A.3)}$$

It should be noted that because of the workers' first order condition of maximizing marginal cost and marginal benefit, we have $\frac{\partial C_i}{\partial e_i} + \frac{\partial P_i}{\partial e_i} = \beta(1+\lambda^K\bar{a}_{\sim i})$ and hence can rewrite (A.3) as

$$\sum_i \frac{\partial e_i^*}{\partial\beta}(1+\lambda^K\bar{a}_{\sim i}) - \beta\sum_i \frac{\partial e_i^*}{\partial\beta}(1+\lambda^K\bar{a}_{\sim i}) - \sum_i \frac{\partial P_i}{\partial \bar{e}_{\sim i}^*}\frac{\partial \bar{e}_{\sim i}^*}{\partial\beta} = 0.$$

Rearranging these elements gives

$$\beta^* = \frac{\sum_i \frac{\partial e_i^*}{\partial \beta}(1+\lambda^K \bar{a}_{\sim i}) - \sum_i \frac{\partial P_i}{\partial \bar{e}^*_{\sim i}} \frac{\partial \bar{e}^*_{\sim i}}{\partial \beta}}{\sum_i \frac{\partial e_i^*}{\partial \beta}(1+\lambda^K \bar{a}_{\sim i})}$$

$$= \frac{\sum_i \frac{\partial e_i^*}{\partial \beta}(1+\lambda^K \bar{a}_{\sim i}) - \sum_i \lambda^P (m - e_i^*) \frac{\partial \bar{e}^*_{\sim i}}{\partial \beta}}{\sum_i \frac{\partial e_i^*}{\partial \beta}(1+\lambda^K \bar{a}_{\sim i})}. \tag{A.4}$$

It should also be noted, however, that, because of our assumption that peer pressure causes no extra utility to workers on average (see Appendix A.2), $\frac{1}{N}\sum_i \lambda^P (m - e_i^*) \frac{\partial \bar{e}^*_{\sim i}}{\partial \beta} \geq 0$. Additionally, from both the expression for optimal effort given in Equation A.2 and $\bar{e}^*_{\sim i} = \frac{\beta[2k(N-1)+\lambda^P]+2k(\beta\lambda^K+\lambda^P)a_i+(\beta\lambda^K+\lambda^P)[2k(N-2)+\lambda^P]\bar{a}_{\sim i}}{[2(N-1)k+\lambda^P](2k-\lambda^P)}$, it follows that $\frac{\partial e_i^*}{\partial \beta} > 0$ and $\frac{\partial \bar{e}^*_{\sim i}}{\partial \beta} > 0$. As a result, $\beta^* \leq 1$ for positive values of λ^P. However, if there is no peer pressure, $\lambda^P = 0$, then we get the familiar result for risk-neutral individuals that $\beta^* = 1$. As m reaches its upper bound, however, $\beta^* = 0$ (see Appendix A.1).

Consider the simplifying case in which all workers have equal ability $a_i = \bar{a}$ and hence exert equal optimal effort $e_i^* = \bar{e}^* = \frac{\beta+\beta\lambda^K\bar{a}+\lambda^P\bar{a}}{2k-\lambda^P}$ (see Appendix A.1). In this case the first-order condition (A.3) simplifies to $\sum_i \frac{\partial \bar{e}^*}{\partial \beta}(1+\lambda^K\bar{a}) - \beta\sum_i \frac{\partial \bar{e}^*}{\partial \beta}(1+\lambda^K\bar{a}) - \sum_i \lambda^P(m-\bar{e}^*)\frac{\partial \bar{e}^*}{\partial \beta} = 0, \Leftrightarrow (1+\lambda^K\bar{a}) - \beta(1+\lambda^K\bar{a}) - \lambda^P(m-\bar{e}^*) = 0$, yielding the solution

$\beta = \frac{(\lambda^P)^2\bar{a}+(2k-\lambda^P)\lambda^K\bar{a}-(\lambda^P m-1)(2k-\lambda^P)}{2(k-\lambda^P)(1+\lambda^K\bar{a})}$, which under the second-order condition $-(1+\lambda^K\bar{a}) + \frac{\lambda^P}{2k-\lambda^P}(1+\lambda^K\bar{a}) < 0, \Leftrightarrow k > \lambda^P$, maximizes firm profits.

A.5 The Impact of Peer Ability on Wages

We now show that all terms in (4) are unambiguously positive.

Because of workers' optimizing behavior, the marginal cost of effort in equilibrium, $\left.\frac{\partial[C(e_i)+P(e_i,\bar{y}_{\sim i})]}{\partial e_i}\right|_{\text{optimal}}$, equals the workers' marginal return $\beta^*(1+\lambda^K\bar{a}_{\sim i})$. Given that $\lambda^K \geq 0$, $\bar{a}_{\sim i} > 0$ and β^* bounded between 0 and 1, this outcome is nonnegative.

$\frac{de_i^*}{d\bar{a}_{\sim i}} > 0$ can be seen from equation (A.2).

$\frac{1}{N}\sum_i \frac{\partial P(e_i,\bar{y}_{\sim i})}{\partial \bar{y}_{\sim i}}\Big|_{\text{optimal}} \geq 0$ holds by assumption (see Appendix A.2).

$\frac{d\bar{y}_{\sim i}}{d\bar{a}_{\sim i}}\Big|_{\text{optimal}} = \frac{d(\bar{a}_{\sim i}+\bar{e}^*_{\sim i})}{d\bar{a}_{\sim i}} = 1 + \frac{d\bar{e}^*_{\sim i}}{d\bar{a}_{\sim i}} > 0$ because of $\frac{d\bar{e}^*_{\sim i}}{d\bar{a}_{\sim i}} > 0$, which can be seen from the expression for $\bar{e}^*_{\sim i}$ given in Appendix A.4.

APPENDIX B: **Variation used in the within-peer group estimator**

Denoting peer group size by N_{ojt}, when $\frac{1}{N_{ojt}}\sum_i \bar{a}_{\sim i,ojt} = \frac{1}{N_{ojt}}\sum_i a_i = \bar{a}_{ojt}$, the within-peer group transformation of equation (6), which eliminates the peer group fixed effect, is

$$\ln(w_{iojt}) - \overline{\ln(w)}_{ojt} = (x^T_{iojt} - \bar{x}^T_{ojt})\beta + (a_i - \bar{a}_{ojt}) + \gamma(\bar{a}_{\sim i,ojt} - \bar{a}_{ojt}) + (\varepsilon_{iojt} - \bar{\varepsilon}_{ojt}),$$

which can in turn be transformed into[24]

$$\ln(w_{iojt}) - \overline{\ln(w)}_{ojt} = (x^T_{iojt} - \bar{x}^T_{ojt})\beta + (a_i - \bar{a}_{ojt}) + \gamma\frac{-1}{(N_{ojt}-1)}(a_i - \bar{a}_{ojt}) + (\varepsilon_{iojt} - \bar{\varepsilon}_{ojt})$$

This calculation shows that, in the within-peer-group transformed model, there is a close association between individual ability and average peer ability: for a one-unit change in individual ability relative to the average peer ability $a_i - \bar{a}_{ojt}$, peer quality relative to the average $\bar{a}_{\sim i,ojt} - \bar{a}_{ojt} = \frac{-1}{(N_{ojt}-1)}(a_i - \bar{a}_{ojt})$ changes by a factor of $\frac{-1}{(N_{ojt}-1)}$. This outcome not only reflects the fact that the better individuals within a peer group have worse peers but also shows that the magnitude of the drop in peer quality for each additional unit of individual ability declines with peer group size. Thus, in the within-peer-group transformed model, individual ability $a_i - \bar{a}_{ojt}$ and peer quality $\bar{a}_{\sim i,ojt} - \bar{a}_{ojt}$ only vary independently if there is heterogeneity in the peer group size N_{ojt}. γ is thus identified by an interaction of N_{ojt} and within-transformed individual ability. As a result, the advantage of being able to control for time-variant peer group fixed effects is countered by the disadvantage that only a particular type of variation is used to identify this effect.

APPENDIX C: **Estimation method**

The solution to estimating equation (5) by nonlinear least squares minimizes the following objective function:

[24] Here, we use $\bar{a}_{\sim i,ojt} - \bar{a}_{ojt} = \frac{N_{ojt}\bar{a}_{ojt} - a_i}{(N_{ojt}-1)} - \frac{(N_{ojt}-1)\bar{a}_{ojt}}{(N_{ojt}-1)} = \frac{-1}{(N_{ojt}-1)}(a_i - \bar{a}_{ojt})$.

$$\min_{\beta,\gamma,a_i,\omega_{ot},\delta_{jt}} M = \sum_i \sum_t \left[\ln(w_{iojt}) - x_{iojt}\beta - a_i - \gamma\bar{a}_{\sim i,ojt} - \omega_{ot} - \delta_{jt} - \theta_{oj} - a_i\right]^2 \quad (9)$$

The algorithm proposed by Arcidiacono et al. (2012) first fixes a_i at starting values and then iterates the following steps:

1. Hold a_i and $\bar{a}_{\sim i,ojt}$ at the values from the previous step and obtain the least square estimates of the now linear model.
2. Update the a_is based on the nonlinear least squares objective function M (see equation (9)), where all other coefficients are set to their estimated values from Step 2. Solving $\frac{\partial M}{\partial a_i}$ for a_i yields functions $a_i = f(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_N)$, which are applied to all a_i repeatedly until convergence, which is ensured under the condition that feedback effects are not too strong (i.e., $\gamma < 0.4$).
3. With the newly updated a_i go back to Step 2 until the parameter estimates converge.

Because the linear model to be solved in Step 2 still includes the high-dimensional fixed effects δ_{jt}, ω_{ot}, and θ_{oj}, we employ a variant of the preconditioned conjugate gradient algorithm to solve this step (see Abowd, Kramarz and Margolis, 1999; Abowd, Creecy, and Kramarz, 2002, for details) that is efficient for very large data matrices.[25]

Because the algorithm does not deliver standard errors and the data matrix is too large to be inverted without hitting computer memory restrictions, we compute the standard errors by implementing a wild bootstrapping with clustering on firms (Cameron, Gelbach, and Miller, 2008).[26] For the baseline model, we verified that when using 100 bootstraps, standard errors were very stable after the 30th bootstrap. Due to the time-consuming nature of the estimation, we then generally used 30 bootstraps for each model.

APPENDIX D: **Imputation of censored wage observations**

To impute the top-coded wages, we first define age-education cells based on five age groups (with 10-year intervals) and three education groups (no post-secondary education, vocational degree, college or university degree). Within each of these cells, following Card et al. (2013), we estimate tobit wage equations separately by year while controlling for age; firm size (quadratic, and a dummy for firm size greater than 10); occupation dummies; the focal worker's mean wage and mean censoring indicator (each computed over time but excluding observations from the current time period); the firm's mean wage, mean censoring indicator, mean years of schooling, and mean university degree indicator (each computed at the current time period by excluding the focal worker observations). For workers observed in only one time period, the mean wage and mean censoring indicator are set to sample means, and a dummy variable is included. A wage observation censored at value c is then imputed by the value $X\hat{\beta} + \hat{\sigma}\Phi^{-1}[k + u(1-k)]$, where Φ is the standard normal cdf, u is drawn from a uniform distribution, $k = \Phi[(c - X\hat{\beta})/\hat{\sigma}]$, and $\hat{\beta}$ and $\hat{\sigma}$ are estimates for the coefficients and standard deviation of the error term from the tobit regression.

[25] We implement the estimation in Matlab based on sparse matrix algebra for efficient data manipulation of the large dummy variable matrices.

[26] Rather than using different observations across bootstraps, this method draws a new residual vector at each iteration, which has the advantage of leaving the structure of worker mobility between firms unchanged across the bootstraps, thereby allowing identification of the same set of worker and firm fixed effects in each bootstrap. Another advantage is that this bootstrap is applicable to clusters of different size.

APPENDIX E: **Bias from wrong peer group definitions**

Defining the peer group at the firm-occupation level leads to two possible error types: excluding relevant peers from outside the occupational group, or including irrelevant peers inside the occupational group. Therefore, we now show that omitting relevant peers generally leads to a downward bias of the true peer effect, whereas including irrelevant peers may not cause any bias. This discussion assumes that the individual and average peer abilities are known because if they need to be estimated, an additional bias may arise from a false definition of peer group.

We denote the average quality of individual i's true and observed peer group by $\bar{a}_{it}^{\text{true}}$ and $\bar{a}_{it}^{\text{obs}}$, respectively, dropping the subscripts j and o (which index firms and occupations) for simplicity. We first suppose that the true model of peer effects is $\ln w_{it} = \mu_{it} + \gamma \bar{a}_{it}^{\text{true}} + u_{it}$, where μ_{it} summarizes the control variables and multiple fixed effects included in the baseline or the within-peer group specification. Because the worker's true peer group is unobserved, we instead regress log wages on μ_{it} and on the average quality of the observed peer group, $\bar{a}_{it}^{\text{obs}}$. The coefficient on $\bar{a}_{it}^{\text{obs}}$ then identifies $\gamma \frac{\text{Cov}\left(\widetilde{\bar{a}_{it}^{\text{true}}}, \widetilde{\bar{a}_{it}^{\text{obs.}}}\right)}{V\left(\widetilde{\bar{a}_{it}^{\text{obs.}}}\right)}$, where $\widetilde{\bar{a}_{it}^{\text{true}}}$ and $\widetilde{\bar{a}_{it}^{\text{obs.}}}$ are residuals from a regression of $\bar{a}_{it}^{\text{true}}$ and $\bar{a}_{it}^{\text{obs.}}$ on the control variables μ_{it}. The relative bias arising from a false peer group definition thus corresponds to the coefficient from a regression of the (residual) average quality of the true peer group on the (residual) average quality of the observed peer group.

Omitting Relevant Peers

If workers interact with all coworkers in their occupation, and also with some coworkers outside their occupation, the average quality of the true peer group is an average of the quality of the observed peer group and the quality of the omitted relevant peers, weighted by group sizes (which for simplicity we assume do not vary across workers and time): $\bar{a}_{it}^{\text{true}} = \frac{N_{\text{obs.}}}{N_{\text{true}}} \bar{a}_{it}^{\text{obs.}} + \frac{N_{\text{omitted}}}{N_{\text{true}}} \bar{a}_{it}^{\text{omitted}}$ which, since the "residual marker" is a linear function, implies $\widetilde{\bar{a}_{it}^{\text{true}}} = \frac{N_{\text{obs.}}}{N_{\text{true}}} \widetilde{\bar{a}_{it}^{\text{obs.}}} + \frac{N_{\text{omitted}}}{N_{\text{true}}} \widetilde{\bar{a}_{it}^{\text{omitted}}}$. Hence, the relative bias becomes

$$\frac{\text{Cov}\left(\widetilde{\bar{a}_{it}^{\text{true}}}, \widetilde{\bar{a}_{it}^{\text{obs.}}}\right)}{V\left(\widetilde{\bar{a}_{it}^{\text{obs.}}}\right)} = \frac{N_{\text{obs.}} + N_{\text{omitted}} \dfrac{\text{Cov}\left(\widetilde{\bar{a}_{it}^{\text{omitted}}}, \widetilde{\bar{a}_{it}^{\text{obs.}}}\right)}{V\left(\widetilde{\bar{a}_{it}^{\text{obs.}}}\right)}}{N_{\text{true}}}$$

Here, if the (residual) qualities of omitted and observed peers are uncorrelated (i.e., $\text{Cov}\left(\widetilde{\bar{a}_{it}^{\text{omitted}}}, \widetilde{\bar{a}_{it}^{\text{obs.}}}\right) = 0$) and the residual peer quality of observed peers does not help predict the residual quality of the relevant omitted peers, omitting relevant peers outside the occupation induces a downward relative bias of $\frac{N_{\text{obs.}}}{N_{\text{true}}}$, which corresponds to the fraction of included relevant peers. This bias is mitigated if the residual qualities of observed and omitted peers are positively correlated. Unconditional on firm-year and firm-occupation fixed effects, $\text{Cov}\left(\widetilde{\bar{a}_{\iota t}^{\text{omitted}}}, \widetilde{\bar{a}_{\iota t}^{\text{obs.}}}\right)$ >0 seems more likely in that more able workers are likely to cluster in the same firm, albeit in different occupations. Conditional on firm-year and firm-occupation fixed effects, however, $\text{Cov}\left(\widetilde{\bar{a}_{\iota t}^{\text{omitted}}}, \widetilde{\bar{a}_{\iota t}^{\text{obs.}}}\right)$ is likely to be closer to zero because the basic sorting of workers into firms and occupations has been controlled for. Hence, the downward bias from omitting relevant peers is likely to be weaker in a regression that controls only for observable characteristics and worker fixed effects than in a regression that also controls for firm-year and firm-occupation fixed effects.

Adding Irrelevant Peers

Next, we suppose that workers interact with no coworkers outside their occupation and only with some workers inside their occupation. In this case, the average quality of the observed peer group is an average of the quality of the true peer group and the quality of the irrelevant included peers, weighted by group sizes (which we again assume to be invariant across workers and time): $\bar{a}_{it}^{\text{obs.}} = \frac{N_{\text{true}}}{N_{\text{obs.}}} \bar{a}_{it}^{\text{true}} + \frac{N_{\text{irrel}}}{N_{\text{obs.}}} \bar{a}_{it}^{\text{irrel}}$, which once again implies $\widetilde{\bar{a}_{\iota t}^{\text{obs.}}} = \frac{N_{\text{true}}}{N_{\text{obs.}}} \widetilde{\bar{a}_{\iota t}^{\text{true}}} + \frac{N_{\text{irrel}}}{N_{\text{obs.}}} \widetilde{\bar{a}_{\iota t}^{\text{irrel}}}$. Hence, the relative bias becomes

$$\frac{\text{Cov}\left(\widetilde{\bar{a}_{\iota t}^{\text{true}}}, \widetilde{\bar{a}_{\iota t}^{\text{obs.}}}\right)}{V\left(\widetilde{\bar{a}_{\iota t}^{\text{obs.}}}\right)} = \frac{\text{Cov}\left(\widetilde{\bar{a}_{\iota t}^{\text{true}}}, \frac{N_{\text{true}}}{N_{\text{obs.}}} \widetilde{\bar{a}_{\iota t}^{\text{true}}} + \frac{N_{\text{irrel}}}{N_{\text{obs.}}} \widetilde{\bar{a}_{\iota t}^{\text{irrel}}}\right)}{V\left(\frac{N_{\text{true}}}{N_{\text{obs.}}} \widetilde{\bar{a}_{\iota t}^{\text{true}}} + \frac{N_{\text{irrel}}}{N_{\text{obs.}}} \widetilde{\bar{a}_{\iota t}^{\text{irrel}}}\right)}$$

If the irrelevant peers are chosen at random conditional on control variables and firm-year and firm-occupation fixed effects (i.e., $\text{Cov}\left(\widetilde{\bar{a}_{\iota t}^{\text{true}}}; \widetilde{\bar{a}_{\iota t}^{\text{irrel}}}\right) = 0$), then this calculation can be simplified to

$$\frac{\text{Cov}\left(\widetilde{\bar{a}_{\iota t}^{\text{true}}}, \widetilde{\bar{a}_{\iota t}^{\text{obs.}}}\right)}{V\left(\widetilde{\bar{a}_{\iota t}^{\text{obs.}}}\right)} = \underbrace{\left(\frac{N_{\text{obs.}}}{N_{\text{true}}}\right)}_{(1)} \underbrace{\frac{\left(\frac{N_{\text{true}}}{N_{\text{obs.}}}\right)^2 V\left(\widetilde{\bar{a}_{\iota t}^{\text{true}}}\right)}{\left(\frac{N_{\text{true}}}{N_{\text{obs.}}}\right)^2 V\left(\widetilde{\bar{a}_{\iota t}^{\text{true}}}\right) + \left(\frac{N_{\text{irrel}}}{N_{\text{obs.}}}\right)^2 V\left(\widetilde{\bar{a}_{\iota t}^{\text{irrel}}}\right)}}_{(2)}$$

The result is two opposing effects: a first term $\left(\frac{N_{\text{obs.}}}{N_{\text{true}}}\right)$ greater than one and a second term smaller than one. The rationale for the first term is that each relevant peer is given weight $\frac{1}{N_{\text{obs.}}}$, instead of the true weight $\frac{1}{N_{\text{true}}}$, thereby reducing the relevant peer's weight in the average peer quality by the factor $\frac{N_{\text{true}}}{N_{\text{obs.}}}$. This reduction in turn increases the coefficient by $\frac{N_{\text{obs.}}}{N_{\text{true}}}$. The second term represents the share of the (weighted) variance of the true peer quality in the overall variance of the observed peer quality, which captures the fact that adding irrelevant peers induces excess variance in observed peer quality, worsening the signal-to-noise ratio. If the (residual) abilities of peers in the true peer group and in the irrelevant peer group are drawn from an i.i.d. distribution with constant variance σ_a^2, then we get $V\left(\widetilde{\bar{a}_{\iota t}^{\text{true}}}\right) = \frac{\sigma_a^2}{N_{\text{true}}}$ and $V\left(\widetilde{\bar{a}_{\iota t}^{\text{irrel}}}\right) = \frac{\sigma_a^2}{N_{\text{irrel}}}$.[27] This latter simplifies the share of variance in the true peer group (the second term in the bias above) to

$$\frac{\left(\frac{N_{\text{true}}}{N_{\text{obs.}}}\right)^2 \frac{\sigma_a^2}{N_{\text{true}}}}{\left(\frac{N_{\text{true}}}{N_{\text{obs.}}}\right)^2 \frac{\sigma_a^2}{N_{\text{true}}} + \left(\frac{N_{\text{irrel}}}{N_{\text{obs.}}}\right)^2 \frac{\sigma_a^2}{N_{\text{irrel}}}} = \frac{\frac{N_{\text{true}}}{N_{\text{obs}}^2}\sigma_a^2}{\frac{N_{\text{true}}}{N_{\text{obs}}^2}\sigma_a^2 + \frac{N_{\text{irrel}}}{N_{\text{obs}}^2}\sigma_a^2} = \frac{N_{\text{true}}}{N_{\text{obs.}}}$$

The two components of the bias therefore cancel each other out leaving no bias in this case; that is, $\frac{\text{Cov}\left(\widetilde{\bar{a}_{\iota t}^{\text{true}}}, \widetilde{\bar{a}_{\iota t}^{\text{obs.}}}\right)}{V\left(\widetilde{\bar{a}_{\iota t}^{\text{obs.}}}\right)} = \left(\frac{N_{\text{obs.}}}{N_{\text{true}}}\right)\left(\frac{N_{\text{true}}}{N_{\text{obs.}}}\right) = 1.$

[27] To illustrate, we group individuals *i* in *M* peer groups (e.g., firms) *j* of size *N* and suppose that individual ability a_{ij} is drawn from a distribution with a zero mean and a covariance structure $\text{Cov}[a_{ij}, a_{i'j}] = \sigma_a^2$ if $i = i'$, and $\text{Cov}[a_{ij}, a_{i'j}] = 0$ otherwise (i.e., there is homoscedasticity and no dependence of error terms of different workers). The variance in average peer ability is then equal to the between-group variance $Var(\bar{a}_j) = Var\left(\frac{1}{N}\sum_{i\epsilon M_j} a_{ij}\right) = \frac{1}{N^2}E\left[\left(\sum_{i\epsilon M_j} a_{ij}\right)^2\right] = \frac{1}{N^2}\frac{1}{M}\sum_j \sum_i \sum_{i'} \text{Cov}[a_{ij}, a_{i'j}] = \frac{1}{N^2}\frac{1}{M}\sum_j N\sigma_a^2 = \frac{\sigma_a^2}{N}$. This inverse proportionality of the variance with respect to group size still holds when heteroscedasticity is allowed between clusters (i.e., , $\text{Cov}[a_{ij}, a_{i'j}] = \sigma_j^2$), which results in $Var(\bar{a}_j) = \frac{\bar{\sigma}^2}{N}$, where $\bar{\sigma}^2 = \frac{1}{M}\sum_j \sigma_j^2$.

Table 1: Skill content, peer group size and wage flexibility for different groups of occupations

	all occupations	5% most repetitive occupations	10% most skilled occupations
Skill Content			
Share without postsecondary education	0.17	0.47	0.04
Share with university degree	0.18	0.01	0.80
To what extent does the following occur in your daily work? (0=never, …, 4=all the time)			
need to learn new tasks and think anew	2.25	1.36	2.98
need to experiment and try out new ideas	1.80	0.96	2.56
need to cooperate with co-workers	2.87	2.80	3.18
Peer Group Size			
median	3	4	3
mean	9.3	12.3	13.1
Wage Flexibility			
St. dev. of log real wage	0.47	0.33	0.37
St. dev. of log real wage residual [a)]	0.38	0.31	0.36
Within-peer group st. dev. of log real wage residual [a)]	0.24	0.20	0.27
Probability of a real wage cut >5%	0.09	0.13	0.04

Note: The table compares all occupations (N=12,832,842) with the 5% most repetitive occupations (N=681,391) and the 10% most skilled occupations (N=1,309,070). See table A1 for a full list of occupations, and section 4.3 of the text for the definition of "repetitive" and "skilled" occupations.
a) Residual from a log-wage regression, after controlling for aggregate time effects, education, and quadratics in firm tenure and age.

Data Source: German Social Security Data, One Large Local Labor Market, 1989-2005.

Table 2: Structure of Sample

No. of workers	2,115,544
No. of firms	89,581
Number of peer groups (occupations within firm-years)	1,387,216
Average number of time periods per worker	6.07
Average number of employers per worker	1.60
Average number of occupations per worker	1.40
Number of peer groups per firm-year	2.30
St. dev. worker fixed effect	0.36
St. dev. average peer fixed effect	0.29
St. dev. change of average peer fixed effect from t-1 to t	0.09
St. dev. change of average peer fixed effect from t-1 to t - Stayers	0.06
St. dev. change of average peer fixed effect from t-1 to t - Movers	0.20
Share of worker-year observations in peer groups without turnover	0.10
Average share of workers replaced by turnover	0.20

Note: The table shows descriptive statistics describing the panel structure of the data set, as well as the variation in wages, peer quality and worker turnover, which we exploit for our subsequent estimates. N=12,832,842.

Data Source: German Social Security Data, One Large Local Labor Market, 1989-2005.

Table 3: Peer Effects in the Full Sample

	(1) outside option only	(2) plus firm fixed effects	(3) plus firm-year fixed effects	(4) plus firm-occupation fixed effects and firm-year fixed effects
Average peer fixed effect	0.414	0.148	0.127	0.011
	(0.002)	(0.002)	(0.001)	(0.001)
Occupation X Year Effects	Yes	Yes	Yes	Yes
Occupation X Firm Effects	-	-	-	Yes
Firm Effects	-	Yes	-	-
Firm X Year Effects	-	-	Yes	Yes

Note: The table shows the effect of average peer quality on log wages. Peer quality is measured as the average fixed worker effect of the co-workers in the same 3-digit occupation at the same firm in the same point of time. In column (1), we only control for worker fixed effects, occupation-by-year fixed effects, and quadratics in age and firm tenure. We then successively add firm fixed effects (column (2)), firm-by-year fixed effects (column (3)), and firm-by-occupation fixed effects (column (4) to control for shared background characteristics. Specification (4) is the baseline specification described in equation (5) in the text. Coefficients can approximately be interpreted as elasticities, and the coefficient of 0.011 in the baseline specification in column (4) implies that a 10% increase in average peer quality increases wages by 0.1%.

Bootstrapped standard errors with clustering at firm level in parentheses. N=12,832,842.

Data Source: German Social Security Data, One Large Local Labor Market, 1989-2005.

Table 4: Peer Effects in Sub-Samples of Occcupations

	(1)	(2)	(3)
Panel A: Peer Effects for Sub-Samples of Low Skilled Occupations			
	5% most repetitive occupations	As in case studies	Low learning content
Average peer fixed effect	0.064	0.067	0.052
	(0.0070)	(0.0116)	(0.0031)
Panel B: Peer Effects for Sub-Samples of High Skilled Occupations			
	10% most skilled occupations	10% most innovative occupations	High learning content
Average peer fixed effect	0.013	0.007	0.017
	(0.0039)	(0.0044)	(0.0028)

Note: The table shows the effect of average peer quality on log wages for different occupational groups. Peer quality is measured by the average worker fixed effect of the co-workers in the same 3-digit occupation at the same firm in the same point of time. All models implement the baseline specification, see equation (5) of the text, and include occupation-by-year, occupation-by-firm, firm-by-year and worker fixed effects, and controls for quadratics in age and firm tenure.

In panel A, column (1), we show the effect for the 5% most repetitive occupations. In panel A, column (2), we show the effect for agricultural helpers, cashiers and data entry workers, which have been used in related case-studies on peer effects in the workplace. In panel A, column (3), we report the effect for occupations characterized by standardized tasks (as the 5% most repetitive occupations) and limited learning content (i.e., cashiers, warehouse workers, drivers, removal workers, cleaners, agricultural helpers, and waiters). In panel B, column (1) we present results for the 10% most skilled occupations, as measured by the share of workers with a college degree in that occupation. In panel B, column (2) we present results for the 10% most innovative occupations, defined by occupation averages of workers' responses to an index of how frequently they need to experiment with new ideas. In panel B, column (3) we present results for occupations with complex tasks and a high learning content (such as doctors, lawyers, scientists, teachers, and academics). See Table A1 for a full list of occupations in each of the sub-samples used in this table, and section 4.3 in the text for an explanation of the way in which the different sub-samples were constructed.

Bootstrapped standard errors with clustering at firm level in parentheses. N=12,832,842.

Data Source: German Social Security Data, One Large Local Labor Market, 1989-2005.

Table 5: Robustness Checks

	(1)	(2)	(3)	(4)
	Within peer group estimator	Wage not imputed	Peer group defined at 2-digit occupational level	Varying coeff. on observables
Panel A: 5% Most Repetitive Occupations				
Average peer fixed effect	0.061	0.086	0.067	0.082
	(0.006)	(0.007)	(0.005)	(0.008)
Panel B: 10% Most Skilled Occupations				
Average peer fixed effect	0.016	0.017	0.010	0.007
	(0.004)	(0.007)	(0.003)	(0.004)

Note: Note: The table reports a number of robustness checks for the effect of average peer quality on log wages. Panel A shows the robustness checks for the group of the 5% most repetitive occupations, as in column (1), Panel A of Table 4. Panel B reports the robustness checks for the group of the 10% most skilled occupations, as in column (1), Panel B, of Table 4.

In column (1), we present the within peer group estimate, see equation (6) in the text. The within-estimator is based on pre-estimated worker fixed effects from the baseline model in equation (5) in the text. The remaining columns refer to our baseline specification given by equation (5) in the text. In column (2), we do not impute censored wage observations. In column (3), we define the peer group at the 2-digit occupational level instead of the 3-digit occupational level. In column (4), we allow the coefficients on the observable characteristics (quadratics in age and firm tenure) to vary between the 5% most repetitive (or 10% most skilled) occupations and the remaining occupations.

Bootstrapped standard errors with clustering at firm level in parentheses. N=12,832,842.

Data Source: German Social Security Data, One Large Local Labor Market, 1989-2005.

Table 6: Peer Pressure or Other Channels? (5% Most Repetitive Occupations)

Panel A: Knowledge Spillovers or Peer Pressure? Age and Tenure Interactions		
	(1)	(2)
	age interaction	tenure interaction
Average peer fixed effect	0.089	0.058
	(0.0088)	(0.0067)
Average peer fixed effect * Age > 35	-0.039	
	(0.0110)	
Average peer fixed effect * Peer tenure>=2		0.007
		(0.0092)
Panel B: Knowledge Spillovers or Peer Pressure? Lagged Peer Group (pre-estimated)		
	inclusion of lagged peer group	
Average peer fixed effect, t	0.054	
	(0.0034)	
Average peer fixed effect, t-1	-0.005	
	(0.0034)	
Panel C: Comlementarities or Peer Pressure?		
	cooperation interaction	
Average peer fixed effect	0.081	
	(0.0078)	
Average peer fixed effect * Co-operation not important	-0.040	
	(0.0110)	
Panel D: Distinguishing between Top vs. Bottom Peers (pre-estimated)		
	top vs. bottom peers	
Average fixed effect of middle 80% peers	0.072	
	(0.0032)	
Average fixed effect of top 10% of peers	0.003	
	(0.0013)	
Average fixed effect of bottom 10% of peers	0.007	
	(0.0014)	

Note: The table provides evidence that the peer effect in wages for the 5% most repetitive occupations are not (only) driven by knowledge spillovers or complementarities in production. In Panel A, we allow the effect of average peer quality on log wages to differ between workers below and above age 35 (column (1)), and between workers who have been in the peer group more and less than 2 years (column (2)). In Panel B, we add the average fixed effects of the lagged peer group to equation (5), based on pre-estimated fixed effects from the baseline model. In Panel C, we allow the peer effect to vary in occupations with below- or above-median shares of workers who report that cooperation is important in their job (column (1)). In Panel D, we split the worker's peers up into the middle 80%, top 10% and bottom 10% according to their ability ranking. This specification is again based on pre-estimated worker fixed effects, and the regression in Panel D is run on the sample for the middle 80% of workers only.

All specifications refer to our baseline specification and cotrol for quadratics in age and firm tenure, worker fixed effects, occuption-by-time fixed effects, firm-by-year fixed effects and firm-by-occupation fixed effects. Bootstrapped standard errors with clustering at firm level in parentheses.

Data Source: German Social Security Data, One Large Local Labor Market, 1989-2005.

Table 7: Heterogeneous Peer Effects (5% Most Repetitive Occupations)

Panel A: Symmetry of Peer Effects (First Differences, Peer Group Stayers, Pre-Estimated Effects)

	Negative Change	Positive Change
Change in average peer fixed effect	0.054	0.047
	(0.008)	(0.008)

Panel B: Heterogeous Effects by Relative Position within the Peer Group

	focal worker below median	focal worker above median
Average peer fixed effect	0.066	0.032
	(0.006)	(0.006)

Panel C: Heterogenous Effects by Gender

	men	women
Average peer fixed effect	0.054	0.075
	(0.008)	(0.008)

Note: In Panel A we investigate whether improvements an deteriorations in average peer quality have similarly sized effects. To do this, we adopt an approach similar to Mas and Moretti (2009) and regress, for peer group stayers, the change in log wages on the change in peer group quality (using the pre-estimated worker fixed effects from our baseline specification), and allow this effect to vary according to whether peer group quality improved or deteriorated. In Panel B we let the peer effect vary by whether the focal worker is above or below the peer-group mean of ability. In Panel C we let the peer effect vary by the focal worker's gender.

Data Source: German Social Security Data, One Large Local Labor Market, 1989-2005.

Table A1: List of Occupations for different Sub-Samples

(1)

5% most repetitive occupations	Share in %
Unskilled laborer, helper (no further specification)	15.12
Packagers, goods receivers, despatchers	11.58
Metal workers (no further specification)	10.66
Postal deliverers	7.58
Assemblers (no further specification)	5.47
Street cleaners, refuse disposers	4.70
Assemblers of electrical parts or appliances	4.68
Cashiers	**4.00**
Railway controllers and conductors	3.96
Laundry workers, pressers	3.69
Machinery or container cleaners and related occupations	2.87
Railway engine drivers	2.80
Milk and fat processing operatives	2.62
Vehicle cleaners, servicers	2.57
Clothing sewers	2.02
Wood preparers	1.96
Metal grinders	1.92
Ceramics workers	1.20
Brick or concrete block makers	1.06
Tobacco goods makers	0.97
Sheet metal pressers, drawers, stampers	0.86
Solderers	0.86
Agricutlural helpers	**0.75**
Model or form carpenters	0.68
Sewers	0.66
Meat and sausage makers	0.61
Stoneware and earthenware makers	0.49
Enamellers, zinc platers and other metal surface finishers	0.42
Leather clothing makers and other leather processing operatives	0.33
Metal moulders (non-cutting deformation)	0.27
Rubber makers and processors	0.27
Other wood and sports equipment makers	0.24
Earth, gravel, sand quarriers	0.22
Machined goods makers	0.20
Moulders, coremakers	0.19
Vulcanisers	0.18
Textile finishers	0.16
Footwear makers	0.15
Other textile processing operatives	0.15
Ready-meal, fruit and vegetable preservers and preparers	0.13
Weavers	0.13
Spinners, fibre preparers	0.12
Textile dyers	0.09
Planers	0.07
Spoolers, twisters, ropemakers	0.05
Post masters	0.05
Radio operators	0.04
Hat and cap makers	0.04
Ship deckhand	0.04
Cartwrights, wheelwrights, coopers	0.03
Rollers	0.03
Wood moulders and related occupations	0.02
Fine leather goods makers	0.02
Fish processing operatives	0.01
Metal drawers	0.01
Jewel preparers	0.01

(2)

Hand-picked occupations with low learning content	Share in %
Salespersons	24.0
Motor vehicle drivers	19.4
Store and warehouse workers	10.9
Household cleaners	8.9
Waiters, stewards	8.0
Unskilled laborer, helper (no further specification)	5.9
Packagers, goods receivers, despatchers	4.5
Gardeners, garden workers	3.7
Goods examiners, sorters, n.e.c.	3.3
Street cleaners, refuse disposers	1.8
Cashiers	1.6
Glass, buildings cleaners	1.5
Laundry workers, pressers	1.4
Transportation equipment drivers	1.4
Vehicle cleaners, servicers	1.0
Earthmoving plant drivers	0.8
Construction machine attendants	0.7
Crane drivers	0.4
Stowers, furniture packers	0.3
Agricutlural helpers	0.3

(3)

10% most skilled occupations	Share in %
Electrical engineers	24.3
Mechanical, motor engineers	13.1
Management consultants, organisors	10.3
Other engineers	7.6
Architects, civil engineers	7.3
Physicians	5.2
Economic and social scientists, statisticians	4.7
Scientists	3.8
Ministers of religion	3.2
Other manufacturing engineers	2.9
Senior government officials	2.9
Physicists, physics engineers, mathematicians	2.5
Technical, vocational, factory instructors	2.0
Legal representatives, advisors	1.9
Primary, secondary (basic), special school teachers	1.7
Chemists, chemical engineers	1.5
University teachers, lecturers at higher technical schools and academies	1.3
Gymnasium teachers	1.3
Pharmacists	0.8
Academics / Researchers in the Humanities	0.8
Garden architects, garden managers	0.4
Survey engineers	0.3
Veterinary surgeons	0.1
Mining, metallurgy, foundry engineers	0.1
Dentists	0.1

(4)

10% most innovative occupations	Share in %
Data processing specialists	38.7
Electrical engineers	24.7
Mechanical, motor engineers	13.3
Architects, civil engineers	7.5
Scientists	3.8
Other manufacturing engineers	2.9
Physicists, physics engineers, mathematicians	2.5
Chemists, chemical engineers	1.5
University teachers, lecturers at higher technical schools	1.3
Musicians	0.9
Interior, exhibition designers, window dressers	0.8
Packaging makers	0.5
Garden architects, garden managers	0.4
Brokers, property managers	0.3
Survey engineers	0.3
Scenery, sign painters	0.2
Veterinary surgeons	0.2
Mining, metallurgy, foundry engineers	0.1
Forestry managers, foresters, hunters	0.1
Coachmen	0.0

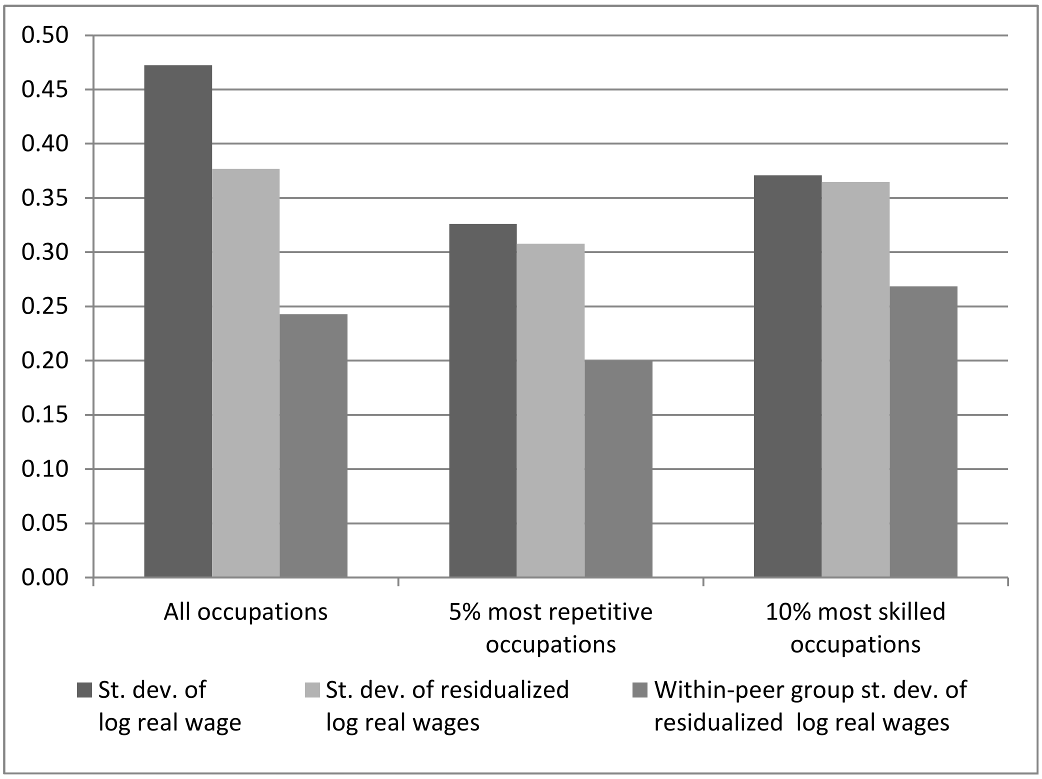
(5)

Hand-picked occupations with high learning content	Share in %
Electrical engineers	20.6
Entrepreneurs, managing directors, divisional managers	18.1
Mechanical, motor engineers	11.1
Management consultants, organisors	8.7
Other engineers	6.5
Architects, civil engineers	6.2
Chartered accountants, tax advisers	5.5
Physicians	4.4
Economic and social scientists, statisticians	4.0
Scientists	3.2
Other manufacturing engineers	2.5
Senior government officials	2.4
Physicists, physics engineers, mathematicians	2.1
Legal representatives, advisors	1.6
Chemists, chemical engineers	1.3
Humanities specialists	0.7
Association leaders, officials	0.6
Survey engineers	0.3
Veterinary surgeons	0.1
Mining, metallurgy, foundry engineers	0.1
Dentists	0.1

Note: The table presents the lists of occupations in for the different sub-samples of occupations used in table 4.

Data Source: German Social Security Data, One Large Local Labor Market, 1989-2005. N=12,832,842.

Figure 1: Variablity of Wages Across and Within Peer Groups.

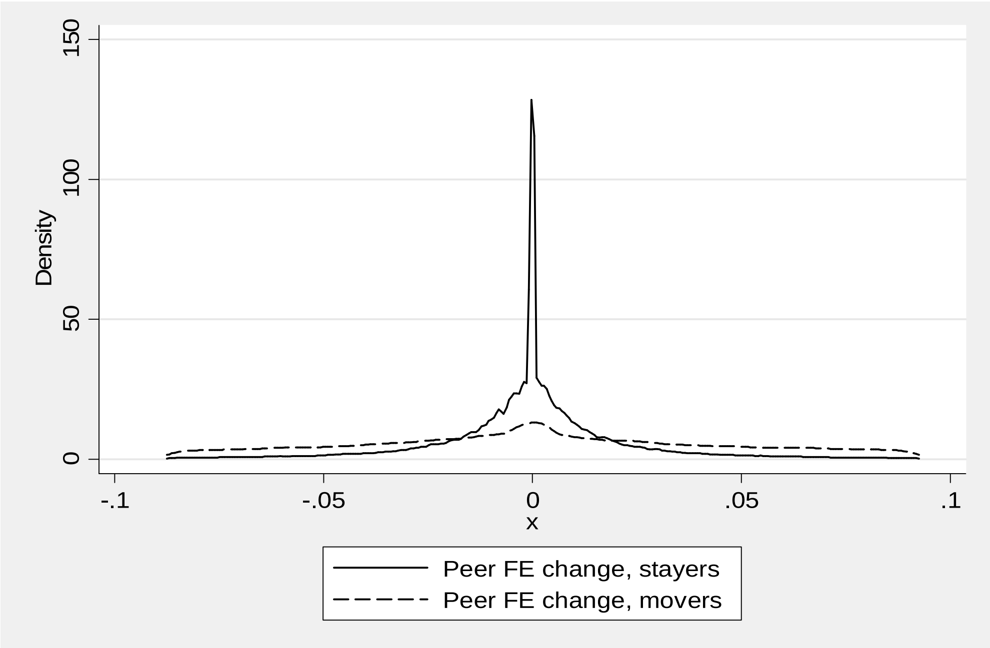


Note: The figure compares all occupations (N=12,832,842), the 5% most repetitive occupations (N=681,391), and the 10% most skilled occupations (N=1,309,070) in terms of the variability of wages.

Residualized wages are computed from a log-wage regression controlling for aggregate time effects, education, and quadratics in firm tenure and age.

Data Source: German Social Security Data, One Large Local Labor Market, 1989-2005.

Figure 2: Variation of the Change in Peer Quality



Note: The figure plots a kernel density estimate of the change in the average peer fixed effect (FE) separately for peer group stayers and peer group movers. Peer group quality varies more strongly for movers. For stayers, there is a mass point at zero, corresponding to stayers in peer groups that had no turnover. The figure is trimmed at the 5% and 95% percentile of the distribution.

Data Source: German Social Security Data, One Large Local Labor Market, 1989-2005.