A Work Project, presented as part of the requirements for the Award of a Master's degree in Economics from the Nova School of Business and Economics.

DYNAMIC PEER EFFECTS IN THE WORKPLACE

DANIEL LUZIO CASTELO

Work project carried out under the supervision of:

Professor Pedro Portugal

10-01-2024

DYNAMIC PEER EFFECTS IN THE WORKPLACE*

By DANIEL CASTELO†

Using exhaustive data on the Portuguese labour force, we present compelling evidence on peer effects as an explanatory variable of workers' future outcomes. Employing a tight identification strategy that circumvents well-known endogeneity concerns, we estimate that a one standard deviation increase in average coworker quality increases workers' next year wages by 4.8%. At the same time, we also observe that despite fading away with time, these peer effects are persistent and remain significant after 10 years. Furthermore, we analyse the heterogeneous effects of peers obtaining results suggestive of ability-related segregation within peer groups as well as ability-specific peer effects persistence.

JEL CODES. C23, J24, J31, L14, M52.

KEYWORDS. Peer Effects, Knowledge Spillovers, Peer Pressure, Fixed Effects, Wage Dynamics.

*I thank my advisor, Professor Pedro Portugal, for his guidance, suggestions, and insightful advice. I thank my friends and family for their unwavering support. I acknowledge the financial support of Fundação para a Ciência e a Tecnologia through the project number PTDC/EGE-ECO/7493/2020.

†Nova School of Business and Economics, Universidade Nova de Lisboa. E-mail: daniel.castelo@novasbe.pt.

I. Introduction

Social interactions play a pivotal role in shaping human behaviour. These influence individuals in ways that extend beyond mere coexistence within a specific space, time, or environment. Such exchanges represent opportunities for the propagation of ideas, behaviours, or social norms among peers, in settings as diverse as workplaces, classrooms, or neighbourhoods.

In this study, we concentrate on the effects of peers in the workplace. Here, literature tends to agree on two main mechanisms through which peer effects can materialize: knowledge spillovers and peer pressure. Knowledge spillovers refer to the knowledge diffusion that flows from on-the-job social interactions. Meanwhile, peer pressure refers to the adoption of certain behaviours by an individual in response to the observation of peer behaviour to avoid feelings of guilt or shame.

We combine the use of uncommonly rich longitudinal data with a particularly tight empirical strategy, incorporating the methodology by Arcidiacono et al. (2012) to the canonical AKM model. With this, we avoid well-known identification concerns (such as the reflection problem or sorting) that often plague similar studies. Also, in line with analyses using comparable data, we focus on the impact of peers on wages rather than on productivity. For this, we use a measure of wages whose evolution proves it to be flexible enough to incorporate peer-induced productivity changes.

Our results present convincing evidence on the existence of peer effects in the workplace. More precisely, we estimate that a one standard deviation increase in contemporaneous average peer quality increases workers' one-year-from-now wages by 4.8%. We also observe a remarkable persistence in the effect on wages of increasing current coworker quality. So much so, that 10 years after exposure (to a one standard deviation increase in average peer quality) the peer effect component is still significant (1.5%). Such persistence supports the notion of knowledge spillovers.

Finally, we also allow for heterogeneity in the response to peers as well as heterogeneity of peer influence. From this new set of specifications, we obtain results suggestive of two different dynamics happening within peer group interactions. First, there appears to exist some kind of ability-related segregation among coworkers - high-ability peers are more influential but also more sensitive to increases in the average quality of fellow high-ability coworkers. A similar dynamic applies to low-ability peers. Simultaneously, the effect of increasing the average quality of high-ability peers seems to be more persistent than the effect of increasing that of low-ability coworkers.

II. LITERATURE REVIEW

Interest in quantifying the effects of colleagues on one's behaviour and outcomes first arose in the field of Education research, where the literature is, as of today, quite extensive.[1] However, the investigation on peer effects was never restricted to the classroom, with some strands of empirical research focusing on other settings such as neighbourhoods (Oreopoulos 2003; Chetty and Hendren 2018), prisons and correctional services (Glaeser, Sacerdote, and Scheinkman 1996; Bayer, Hjalmarsson, and Pozen 2009), among many others. Nevertheless, and despite these being already theorized as far back as the 1800s, the empirical study of peer effects applied to the workplace is a rather recent innovation, resulting in a literature that is still quite scarce.

Marshall (1890) was the first to hypothesize that social interactions happening between coworkers create learning opportunities that will, ultimately, enhance productivity. This idea of interactions as a source of overall income growth has also been adopted by influential macroeconomic growth models. Following Marshall's hypothesis, and building on the works by Solow and Deninson, Lucas (1988) proposes a model for long-run economic growth that includes the accumulation of human capital as an explaining variable. The idea is that knowledge diffusion among knowledge producers will have strong positive spillovers across coworkers. Similarly, in his theory on endogenous technological growth, Romer (1990) also builds on this idea of human capital spillovers.

Now, concentrating on the empirical literature, one may divide it into *laboratory experiments* and *studies using real-world data*. On the side of experimental studies, Falk and Ichino (2006) studied a group of high school students who were recruited from a city in the canton of Zurich.

1. See Sacerdote (2011) for an overview.

These students were tasked with stuffing letters into envelopes and were randomly divided into two treatments: they would either sit alone or sit in pairs while performing the assigned task. Still, on the experimental side of research, Guryan, Kroft, and Notowidigdo (2009) resort to a quasi-experiment to evaluate the effect of peers. More specifically, they use random assignment in golf tournaments to study if a playing partner's ability affects one's performance. While the first study found evidence for peer effects, the second did not.

As for the observational studies, these often come from very specific settings (usually coming from one sole firm or occupation) with very limited external validity. For example, Mas and Moretti (2009) focus on cashiers from one large supermarket chain and verify that workers' productivity increases when they work with more productive coworkers. Similarly, Bandiera, Barankay, and Rasul (2010) concentrate on soft-fruit pickers in a large UK farm, finding that workers are more productive when working with a more productive friend. As for studies focusing on single occupations, Jackson and Bruegmann (2009) look at a longitudinal dataset of student scores linked to teachers in North Carolina. In this study, the authors obtain evidence for peer effects among teachers, with historical peer quality explaining 20 percent of the own-teacher effect.

Recently, some literature started to be developed using more comprehensive and representative datasets. Here, we can mention the work developed by Cornelissen, Dustmann, and Schönberg (2017) and by Jarosch, Oberfield, and Rossi-Hansberg (2021). These studies use German social security records, containing information for the entire workforce of a sample of establishments. Similarly, Battisti (2017), and Hong and Lattanzio (2022) exploit a dataset of social security administrative data with detailed information on the labour force and firms of the Italian region of Veneto.[2] In addition to these, it is also worth mentioning the work carried out by Portugal et al. (2022) who analysed a dataset containing information on a broad segment of the Portuguese working population. Moreover, unlike the preceding studies that focus on the impact of peers on *productivity*, this strand of literature focuses on the impact of peers on *wages*.

2. Containing, virtually, all the employees working in the Private Sector.

The studies mentioned in the above paragraph can also be divided according to their approach regarding the evaluation of peer effects. Whereas Battisti (2017), Cornelissen, Dustmann, and Schönberg (2017), and Portugal et al. (2022) focus on contemporaneous peer effects (i.e., the effects of peers within the same period); Jarosch, Oberfield, and Rossi-Hansberg (2021), and Hong and Lattanzio (2022) focus on dynamic effects (i.e., the effect of peers across time). All these studies found evidence of positively significant peer effects both on current and future wages. The only exception is the study by Cornelissen, Dustmann, and Schönberg which only found small peer effects on wages, on average for all occupations.[3]

III. THEORETICAL FRAMEWORK

III.A. Identification Strategy

Among the empirical studies working with non-experimental data, the literature may be divided into two main strands according to the way in which studies attempt to overcome the identification challenges associated with peer effects quantification. The first strand uses policy changes as a source of exogenous variation in peer outcomes. For example, some studies use the dismissal of Jewish and "politically unreliable" scientists in Nazi Germany as an exogenous shock, whereas others use the emigration of Soviet mathematicians upon the dissolution of the Soviet Union (Waldinger 2012; Borjas and Doran 2015). As for the second strand, it explores time variation in outcomes and changes in peer group formations via longitudinal datasets in order to control for unobserved characteristics and for endogenous selection into groups (Cornelissen, Dustmann, and Schönberg 2017). Our study falls under the second category.

When working with observational data, there are significant challenges to the identification of peer effects, as extensively discussed by Manski (1993), as well as Angrist (2014). On this note, we can highlight three main contentious issues to our empirical analysis: (i) the *reflection problem*, (ii) *sorting and homophily*, and (iii) *unobserved correlated shocks*. The reflection problem surges

3. Even though, they obtain significant coefficients when they focus on low-skilled occupations.

from the troubling nature of outcome-on-outcome peer models, where peer effects are identified via regressing individuals' outcomes on the mean outcome of the relevant reference group, with or without controlling for individual or group mean characteristics (e.g., regressing each individual's hourly wage on its peer group's average hourly wage). Sorting and homophily refer to the non-random distribution of workers across firms, occupations, or peer groups. The potential presence of correlated shocks refers to the possible existence of unobserved time-varying peer group-specific shocks in wages that are correlated with changes in average peer group quality.

To circumvent the reflection problem, we will regress future wages on a predetermined variable – average individual fixed effects of coworkers. This way, we avoid the concerns that plague outcome-on-outcome models for peer effects. Moreover, to estimate the aforementioned fixed effects, we will be resorting to the framework developed by Arcidiacono et al. (2012). Through an iterative algorithm, this method allows the estimation of both individual and coworker average fixed effects, where the latter is a linear combination of a worker's peers' individual fixed effects.

To tackle sorting, homophily and the possibility of unobserved correlated shocks, we build on the firm and worker effects model by Abowd, Kramarz, and Margolis (1999). As such, we will be including both individual and peer group-year fixed effects. The latter, essentially, corresponds to establishment-by-job-title-by-year fixed effects. With this, we impose a highly dimensional set of fixed effects that control for the aforementioned sources of endogeneity. Compared to the typical specifications of empirical literature, this identification strategy is a rather tight one.

III.B. Mechanisms

The economic literature, generally, identifies two mechanisms through which peer effects may materialize. On the one hand, the interactions that arise from working in teams enable the comparison between coworkers which may lead workers to feel an urge to be, at least, as good as their peers. This is, commonly, defined as *peer pressure*. On the other hand, the literature also discusses social interactions among coworkers as an opportunity for these to learn with each other, enhancing their productivity. This mechanism is, usually, defined as *knowledge spillovers*. Whether one or the other is the most prevalent is a debated topic in the literature, with no widely accepted answer.

On the side of peer pressure, Mas and Moretti (2009) attribute their results (of positive peer effects on productivity) to social pressure as they only find evidence for coworkers that are visible to each other. They also discuss how this mechanism of social pressure may partially internalize the free-riding problems that may often arise within firms. Similarly, despite obtaining weak evidence for peer effects, on average, Cornelissen, Dustmann, and Schönberg (2017) observe substantial peer effects estimates in specific occupations whose productivity is more visible and that are less knowledge-intensive (e.g., supermarket cashiers, fruit pickers). This result leads the authors to attribute most of the coworker effects to the peer pressure mechanism.

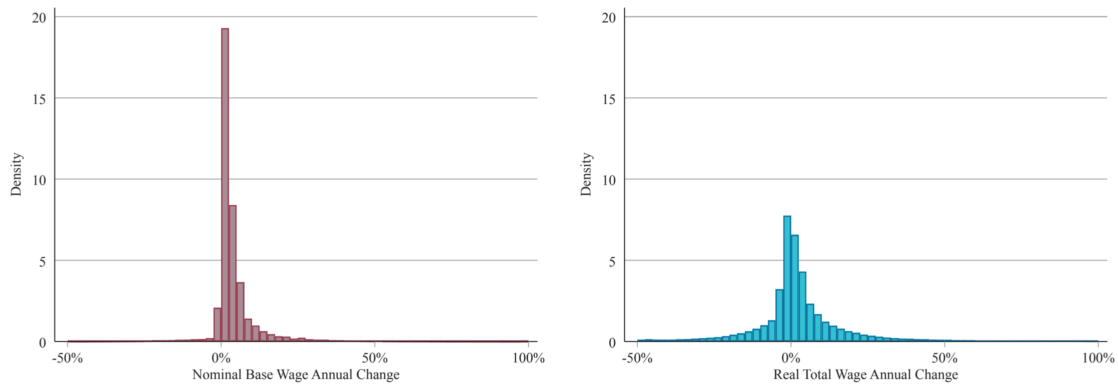
As for knowledge spillovers, Jackson and Bruegmann (2009) conclude that the evidence they obtained for peer effects among teachers, is suggestive of learning among coworkers. This conclusion derives from the fact that less experienced teachers (who are assumed to be the ones more susceptible to learning from coworkers) are the most sensitive to changes in peer quality. Additionally, the study on the dynamics of peer effects by Jarosch, Oberfield, and Rossi-Hansberg (2021) also provides evidence for knowledge spillovers. This is suggested by the fact that workers' future wage growth is much more sensitive to the wages of higher-paid coworkers (who are assumed to be more productive) than to those of lower-paid ones.

III.C. Identification of Peer Effects Through Wages

For the purpose of this study, and in line with the similar studies covered in Section II, we will use wages to capture peer effects. This requires some degree of wage flexibility so that individual wages can internalize productivity changes induced by peer effects. However, the wage setting process in Portugal is distinctively rigid. For example, Article 129 of the Portuguese Labour Code explicitly forbids (nominal) base wage cuts since the 1950s.[4] Such a landscape may critically

4. Although some exceptions are contemplated in the body of law.

FIGURE 1. HISTOGRAM OF THE ANNUAL CHANGE OF DIFFERENT WAGE MEASURES



(A) NOMINAL BASE WAGES (B) REAL TOTAL WAGES

Notes. The figures plot the histograms of yearly growth rates of the two wage measures. In addition to these, we also plot the distributions of two measures' growth rates along the time and mobility dimensions in Figures A.1, A.2, A.3, and A.4. The figure is trimmed at the -50% and 100% thresholds. *Source.* Quadros de Pessoal, 2002-2021.

jeopardize the validity of our study, as wages may not be able to internalize existing peer effects.

Bearing this in mind, we try to avoid the issue of nominal base wage rigidity by including, in our wage variable, all benefits and payments received by workers.[5] This way, we are able to capture productivity bonuses and other additional payments that may give a more revealing insight into productivity changes. In addition to this, we will also be using real wages (deflated to 1986 prices), to account for inflation-led wage reductions. To better understand if this hypothesis holds we build histograms on the annual percentage changes of *nominal base* wages and *real total* wages.

Figure 1 plots a compelling story in favour of our argument. Even though nominal base wages are remarkably rigid, the same cannot be said about real total wages where rigidity seems to be significantly attenuated. Moreover, using the same threshold used by Cornelissen, Dustmann, and Schönberg (2017) to evaluate downward wage flexibility, we obtain that 17% of the worker-year observations in our sample observe, at some point, a wage cut larger than 5% from one year to another (16% for stayers and 29% for movers). This indicates that our measure for wages can

5. More precisely, we include base wages, overtime payments, and regular and irregular benefits. Further, as we work with hourly wages, we divide the previous amount by the number of hours worked.

reasonably incorporate peer-related losses in productivity, even for the case of workers who did not switch jobs. For reference, this value was around 9% for stayers in the study by Cornelissen, Dustmann, and Schönberg. As such, the (real total) wage setting scenario in Portugal should be flexible enough to accommodate productivity changes deriving from peer effects.

IV. METHODOLOGY

IV.A. *Baseline Specification*

As mentioned above, our model makes use of an extended AKM model (Abowd, Kramarz, and Margolis 1999) including peer effects, calculated via the methodology devised by Arcidiacono et al. (2012). We start by implementing the following baseline specification, which we estimate individually for each time horizon h (with $1 \leq h \leq 10$):

$$Y_{i,t+h} = \alpha_i + X_{i,t}^{'}\,\beta + \gamma_h \bar{\alpha}_{\sim i,t} + \delta_{n,t} + \varepsilon_{i,t+h}\ , \tag{1}$$

$$\bar{\alpha}_{\sim i,t} = \frac{\sum_{j \in \mathbb{M}_{n,t}} (\alpha_j) - \alpha_i}{|\mathbb{M}_{n,t}| - 1}$$

$$\text{for } i \in \mathbb{M}_{n,t}$$

where $Y_{i,t+h}$ is the natural logarithm of real hourly wage, in 1986 prices, of worker i, in period $t + h$; $X_{i,t}^{'}$ is a vector of relevant time-varying covariates;[6] $\delta_{n,t}$ refers to the peer-group-by-year fixed effect. The term α_i is the worker i's individual fixed effect, constituting a proxy of a worker's long-term ability. The term $\bar{\alpha}_{\sim i,t}$ refers to the average of worker i's coworkers' individual fixed effects at time t. Here, more specifically, $\mathbb{M}_{n,t}$ represents the set of workers belonging to peer group n at time t (to which worker i belongs to), meaning that $|\mathbb{M}_{n,t}|$ represents this set's cardinality (i.e., the peer group's number of workers in the relevant time period); and $\sum_{j \in \mathbb{M}_{n,t}} (\alpha_j) - \alpha_i$ corresponds to the sum of the individual fixed effects of all workers j belonging to peer group n at time t minus the individual fixed effect of worker i. Similar to the worker fixed effect, this component is a proxy

6. More precisely, age, squared age, tenure, squared tenure, schooling, and squared schooling.

of the average quality of worker i's coworkers at time t. The coefficient γ_h is the parameter of interest. It measures the effect of coworker quality on workers' future wages (h years from now).

IV.B. Estimation

To implement the specification from Equation 1, we exploit the longitudinal nature of our dataset. Only thanks to it are we able to obtain workers' individual fixed effects as well as our proxy for average coworker quality - which we construct via linear combinations of coworkers' individual fixed effects. Note that the use of fixed effects to unveil the impact of peers is especially adequate as the unobserved characteristics that make a ”good peer” tend to be time-invariant. This method is known to yield consistent peer effect estimates (Arcidiacono et al. 2012).

As the unknown peer effect is estimated through a combination of similarly unknown worker fixed effects, we end up with a nonlinear optimization problem, for which least squares estimation is infeasible due to the problem's dimensionality. As such, we must resort to an iterative process to obtain the estimates of the baseline specification. This iterative process alternates between estimating the individual fixed effects and the peer effect parameter, γ_h, reducing the sum of squared errors with each iteration until these estimates converge to an imposed criterion of precision.[7]

IV.C. Identification Challenges

As previously mentioned, in Section III.A, there are three potential sources of endogeneity that may challenge the robustness of our results. These are the reflection problem, sorting and homophily, and the presence of unobserved correlated shocks.

The Reflection Problem. First introduced by Manski (1993), it refers to the problem arising from the use of outcome-on-outcome models to measure peer effects. In other words, due to the simultaneous use of a variable as a regressand (to identify individual behaviour) and as a regressor (to measure mean group behaviour) an identification problem surges as the ”peer effect” is a mechanical phenomenon, with no causality implications. Take the simple (yet extreme) example:

7. To implement this process, we use the *regpeer* ado file developed by Portugal et al. (2022).

$z_{i,j} = \phi + \psi \bar{z}_j + \omega_{i,j}$

where i indexes the individual and j the peer group. The term $z_{i,j}$ is a particular variable that is observed for individual i belonging to peer group j and $\bar{z}_j$ is the peer group mean of that same variable. The parameter of interest is ψ which, in principle, should ”measure” the effect of peer behaviour on one's behaviour. However, due to the nature of OLS estimation, this coefficient will always correspond to the unity (1), independently of the particular variable under analysis.

Again, this is an extreme example and, typically, empirical literature uses more elaborate models where they include additional controls or even use leave-out means in an attempt to avoid this problem. Here, the theory diverges. Manski (1993) believes that the previously described problem is a special case that can be avoided by including enough pre-determined information on how individuals self-select into peer groups. In contrast, Angrist (2014) does not believe this to be a credible solution, as it raises new concerns regarding the econometric validity of the estimates.

We avoid these concerns by simply not using an outcome-on-outcome model. By using the framework established by Arcidiacono et al. (2012) we, essentially, link workers' wages to coworkers' mean individual fixed effects - a long-term ”predetermined” characteristic (that we can calculate thanks to the longitudinal nature of the dataset) - rather than the peer group's average wages.

Sorting and Homophily. Sorting refers to the endogenous distribution of workers across firms, job titles, and peer groups (firm-job title combinations). Homophily refers to the similar behaviour of peers due to shared individual characteristics and shared environment rather than due to peer effects deriving from social interactions among a particular group of coworkers.

As such, to account for the potential sorting of high-quality workers into high-paying firms we shall include firm fixed effects. To control for sorting into high-paying job titles we use job title fixed effects. Finally, to address sorting into high-paying firm-job title combinations we include firm-job title fixed effects. The latter specification of fixed effects, due to its multi-way structure, encompasses both firm and occupation fixed effects as well as any combination of the two. As for homophily, we circumvent the issue of workers having similar ex-ante characteristics (which may be erroneously attributed to peer effects) by including worker fixed effects.

Unobserved Correlated Shocks. It refers to the possible existence of time-varying wage shocks at the peer group level correlated to shocks in peer group quality due to some unobserved shock. Take the illustrative example of a factory that invests in a new technology. Assume this technology is specific to one job title only and that it requires higher-quality workers to operate it. In such a scenario, we can expect a simultaneous increase in wages and worker quality due to the exemplified unobserved shock. The existence of such correlated shocks would lead to a biased γ estimate. To avoid this issue we opted to further extend the already extensive establishment-job title fixed effects to account for the full set of establishment-job title-year effects. This further tightens our identification strategy and, ultimately, enhances the robustness of our results.

IV.D. Heterogeneous Effects Specification

The baseline specification rests on the assumption that all workers *respond* to coworkers and *influence* coworkers similarly. But this assumption is quite restrictive and may not be, particularly, interesting to understand the dynamics at play. That is why we build upon the baseline specification, and allow for heterogeneous peer effects. More precisely, we will look at (i) the effect *on* heterogeneous workers, (ii) the effect *of* heterogeneous peers, and (iii) a combination of both (effect of heterogeneous peers on, equally, heterogeneous workers). To implement this model, we will use the worker fixed effects obtained from the baseline specification.

For the purpose of this exercise, we will consider two subgroups within each peer group: one of *high-ability* peers (denoted by "+") and another of *low-ability* peers (denoted by "-"). To construct each of these subgroups, we use peer group leave-out median of individual fixed effects as a threshold.[8] Workers whose individual effects are above the threshold will the considered high-ability peers and workers below it will be considered low-ability peers.

8. As mentioned before, individual effects can be read as a proxy for innate ability

Heterogeneous Responsiveness to Peers. To allow heterogeneity in the *response* to peers, we will separately implement the following specification, for each ability-specific peer subgroup, $\mathbb{Q}(o)$:

$$Y_{i,t+h} = \hat{\alpha}_i + X_{i,t}' \, \beta + \gamma_{h,\mathbb{Q}(o)} \hat{\alpha}_{\sim i,t} + \hat{\delta}_{n,t} + r_{i,t+h} \tag{2}$$

$$Y_{i,t+h} - \hat{\alpha}_i - \hat{\delta}_{n,t} = X_{i,t}' \, \beta + \gamma_{h,\mathbb{Q}(o)} \hat{\alpha}_{\sim i,t} + r_{i,t+h} \ , \quad i \in \mathbb{M}_{n,t}^{\mathbb{Q}(o)} \tag{3}$$

$$\hat{\alpha}_{\sim i,t} = \frac{\sum_{j \in \mathbb{M}_{n,t}} (\hat{\alpha}_j) - \hat{\alpha}_i}{|\mathbb{M}_{n,t}| - 1}$$

$$\text{for } i \in \mathbb{M}_{n,t}, \ \mathbb{Q}(o) \in \{-,+\} \text{ and } \mathbb{M}_{n,t}^{\mathbb{Q}(o)} \subset \mathbb{M}_{n,t}$$

where $\mathbb{Q}(o)$ indexes the heterogeneous-response peer subgroups we have constructed, meaning that $\mathbb{M}_{n,t}^{\mathbb{Q}(o)}$ represents each set of the two ability-specific peer subgroups - which are both contained in the broader peer group set $\mathbb{M}_{n,t}$. The term $\hat{\alpha}_{\sim i,t}$ refers to the mean individual fixed effects of individual i's coworkers (which we obtain from the baseline). The parameter of interest is $\gamma_{h,\mathbb{Q}(o)}$ and it, separately, measures the effect of coworker quality on future wages of high- and low-ability workers, i.e., the heterogeneous responses to peer effects of each ability-specific subgroup.

Heterogeneous Peer Influence. To allow for heterogeneity in *peer influence*, we need to extend the baseline specification to accommodate for heterogeneous influence:

$$Y_{i,t+h} = \hat{\alpha}_i + X_{i,t}' \, \beta + (\gamma_h^+) \, (\hat{\alpha}_{\sim i,t}^+) + (\gamma_h^-) \, (\hat{\alpha}_{\sim i,t}^-) + \hat{\delta}_{n,t} + r_{i,t+h} \tag{4}$$

$$Y_{i,t+h} - \hat{\alpha}_i - \hat{\delta}_{n,t} = X_{i,t}' \, \beta + (\gamma_h^+) \, (\hat{\alpha}_{\sim i,t}^+) + (\gamma_h^-) \, (\hat{\alpha}_{\sim i,t}^-) + r_{i,t+h} \tag{5}$$

$$\hat{\alpha}_{\sim i,t}^{\mathbb{S}(k)} = \begin{cases} \dfrac{\sum_{j \in \mathbb{M}_{n,t}^{\mathbb{S}(k)}} (\hat{\alpha}_j) - \hat{\alpha}_i}{|\mathbb{M}_{n,t}^{\mathbb{S}(k)}| - 1}, & i \in \mathbb{M}_{n,t}^{\mathbb{S}(k)} \\ \dfrac{\sum_{j \in \mathbb{M}_{n,t}^{\mathbb{S}(k)}} \hat{\alpha}_j}{|\mathbb{M}_{n,t}^{\mathbb{S}(k)}|}, & i \notin \mathbb{M}_{n,t}^{\mathbb{S}(k)} \end{cases}$$

$$\text{for } i \in \mathbb{M}_{n,t}, \ \mathbb{S}(k) \in \{-,+\}, \text{ and } \mathbb{M}_{n,t}^{\mathbb{S}(k)} \subset \mathbb{M}_{n,t}$$

where $\mathbb{S}(k)$ indexes the heterogeneous-influence peer subgroups we have constructed. The term $\hat{\alpha}_{\sim i,t}^{\mathbb{S}(k)}$ refers to the ability-specific mean individual fixed effects of individual i's coworkers. Here, we have two cases: (i) worker i belongs to the particular ability-specific peer subgroup $\mathbb{M}_{n,t}^{\mathbb{S}(k)}$, meaning that $\bar{\alpha}_{\sim i,t}^{\mathbb{S}(k)}$ corresponds to the leave-out mean individual fixed effects of worker i's coworkers belonging to the particular ability-specific peer group; and (ii) worker i does not belong to the particular ability-specific peer subgroup $\mathbb{M}_{n,t}^{\mathbb{S}(k)}$, which means that $\bar{\alpha}_{\sim i,t}^{\mathbb{S}(k)}$ refers to the simple average of individual fixed effects of worker i's coworkers belonging to the particular ability-specific peer subgroup. The parameters of interest are $\gamma_h^{(+)}$ and $\gamma_h^{(-)}$. These, independently, measure the effect of ability-specific coworkers on workers' future wages.[9]

Heterogeneous Responsiveness to Heterogeneous Peer Influence. To allow for heterogeneity in the *response* to heterogeneous *peer influence*, we simply combine the previous specifications:

$$Y_{i,t+h} - \hat{\alpha}_i - \hat{\delta}_{n,t} = X_{i,t}'\,\beta + (\gamma_{h,\mathbb{Q}(o)}^{+})\,(\hat{\alpha}_{\sim i,t}^{+}) + (\gamma_{h,\mathbb{Q}(o)}^{-})\,(\hat{\alpha}_{\sim i,t}^{-}) + r_{i,t+h}\,, \quad i \in \mathbb{M}_{n,t}^{\mathbb{Q}(o)} \tag{6}$$

$$\hat{\alpha}_{\sim i,t}^{\mathbb{S}(k)} = \begin{cases} \dfrac{\sum_{j\in\mathbb{M}_{n,t}^{\mathbb{S}(k)}}(\hat{\alpha}_j) - \hat{\alpha}_i}{|\mathbb{M}_{n,t}^{\mathbb{S}(k)}| - 1}, & i \in \mathbb{M}_{n,t}^{\mathbb{S}(k)} \\ \dfrac{\sum_{j\in\mathbb{M}_{n,t}^{\mathbb{S}(k)}}\hat{\alpha}_j}{|\mathbb{M}_{n,t}^{\mathbb{S}(k)}|}, & i \notin \mathbb{M}_{n,t}^{\mathbb{S}(k)} \end{cases}$$

$$\text{for } i \in \mathbb{M}_{n,t};\ \mathbb{Q}(o),\ \mathbb{S}(k) \in \{-,+\};\ \text{and } \mathbb{M}_{n,t}^{\mathbb{Q}(o)},\ \mathbb{M}_{n,t}^{\mathbb{S}(k)} \subset \mathbb{M}_{n,t}$$

where where $\mathbb{Q}(o)$ indexes the heterogeneous-response peer subgroups we have constructed, and $\mathbb{S}(k)$ indexes the heterogeneous-influence ones. The parameters of interest are $\gamma_{h,\mathbb{Q}(o)}^{+}$ and $\gamma_{h,\mathbb{Q}(o)}^{-}$ and they, separately, measure the effect of ability-specific coworkers on the future wages of ability-specific workers. Put differently, they quantify the effect of high-ability peers on high- and low-ability workers, as well as the effect of low-ability peers on high- and low-ability workers.

V. DATA

Quadros de Pessoal (QP) is a longitudinal employer-employee-matched dataset for Portugal, covering the years between 1986 and 2021.[10] It is compiled annually via a mandatory survey

9. Where γ_h^{+} captures the effect of high-ability peers, and γ_h^{-} the effect of low-ability peers.
10. With the exception of 1990 and 2001.

collected by the Portuguese Ministry of Employment, containing information for all firms with, at least, one wage earner. From 1994 onward, the collected information refers to the month of October. Its compulsory nature ensures that concerns like panel attrition and measurement error are significantly attenuated. In fact, the survey covers all wage earners in Portugal, with the sole exceptions of civil servants and self-employed individuals. Such a level of coverage over the total labour force makes QP a remarkably rich and comprehensive dataset that is unique in the world.

Quadros de Pessoal contains detailed information on workers, reporting demographic variables (such as age, gender, and education) as well as employment-related information (tenure, collective agreement, job title, or wages and other benefits).[11] Moreover, the dataset also reports valuable information on establishments and firms (headcount, turnover, location, economic activity, or legal form). A valuable feature of QP is the fact it enables linkage between workers, establishments, and firms through common identifiers. A simple yet powerful tool that allows us to trace employees' professional history - the firms they have worked for, the jobs they have performed, and more importantly (to our analysis) the coworkers they have encountered.

V.A. Sample Selection

Our analysis exploits a sample that covers the years between 2002 and 2021. This choice was made in order to avoid breaks in the data, flowing from missing years. We restrict the sample to only include workers between the ages of 18 and 64. Only full-time workers with, at least, 120 monthly hours of work are included. We dropped workers whose base wage was below 80% of the national minimum wage. A limit on tenure was also imposed: values above the 50-year threshold were not included. Workers not assigned to any collective bargaining agreement were dropped, as well as those belonging to residual occupation categories (apprentices and similar occupations). In addition to these restrictions, we further exclude each year's top 1% hourly wages - to prevent extreme outliers from affecting our results. We also, naturally, limit peer group sizes so that these

11. Concerning other benefits, besides base wages, Quadros de Pessoal has information on regular and irregular benefits, and overtime payments.

have, at least, two workers. Besides this, we also exclude the top 1% peer group sizes, meaning that we exclude all peer groups with more than 33 workers. This way we avoid including peer groups whose dimension would not allow for interactions between all workers. Finally, as recommended by Abowd, Creecy, and Kramarz (2002), we use the largest connected set of observations.

As we work with different time horizons, the sample size used in our regressions is evermore limited to a subset of the original sample as the time horizon h increases. For instance, when considering time horizon $h = 1$, we can only use the years between 2002 and 2020 (while values for the years between 2003 and 2021 will be used for the dependent variable, future wage in one year). Similarly, when working with $h = 10$ we can only include the years between 2002 and 2011. Considering the overall sample of this study, it is comprised of 18,593,626 worker-year observations on 3,518,197 workers employed in 287,133 firms. Overall these account for a total of 4,446,157 peer group-year observations, which have, on average 4.18 workers. See Table A.3 in the Appendix for detailed descriptive statistics.

V.B. Relevant Definitions

Worker, Establishment, Firm, and Job Title. We define as *worker* each uniquely identified wage earner present in the sample. The term *establishment* refers to any kind of physical location in which workers perform tasks in exchange for a wage. Finally, *firm* describes a legal entity that employs workers. By construction, no wage earner may belong to more than one firm at any given time t. A firm may have several establishments, but an establishment can either be a firm or part of a firm. As for *job title*, we define it as each professional category that falls under the same collective wage bargaining agreement.

Peer Group. In this study, we will define peers as employees who for a given year share the same job title and work in the same establishment of the same firm. This definition of peers is in line with that of Portugal et al. (2022), whose analysis also focuses on Portuguese data. It also constitutes a stricter definition than those used by similar observational studies. This restrictive approach is to ensure that "peers" refer to workers performing similar tasks within the same physical workplace – which we assume to be the closest way to unveil on-job social interactions.

VI. Results

VI.A. Baseline Results

We start by implementing the specification laid out in Equation 1. We report the estimates for our parameter of interest, γ_h, in Table 1, clustering standard errors at the worker level. Each column represents a different time horizon, h, with Panel A reporting the effect of peer quality on wages after 1 to 5 years and Panel B reporting the same effect after 6 to 10 years. The same estimates are also plotted in Figure 1 where we include the 95% confidence intervals.

We obtain clear evidence supporting the idea that peer quality has a significant positive impact on future wages. More precisely, we obtain that a, ceteris paribus, 10 percent increase in contemporaneous average peer quality increases next year's wage by 1.51 percent (1.41 log points).[12] In other words, a one standard deviation increase in the average quality of current coworkers will increase next year's wages by 4.8 percent.[13]

At the same time, the decaying trend in the peer effects coefficient suggests that the returns from increasing contemporaneous coworker quality are each year smaller. This hints that the exposure to better peers happening in the contemporaneous period ($h = 0$) fully materializes in wages within the first year. An effect that, then, gradually fades away as the time horizon increases. However, this effect does not fully disappear within the first 10 years after exposure - as a one standard deviation increase in average peer quality today yields a 1.5% positive impact on 10-years-from-now wages. In fact, after the seventh year (post-exposure) the effect stops fading away and appears to become systematic, possibly suggesting that the effect of being exposed to better peers never fully disappears. This evolution is well captured by Figure 1.

---

12. As stated before, we use the average of coworker individual fixed effects as a proxy for peer quality.
13. (0.151 $\times$ 0.32) Table A.4 reports standard deviations and other statistics for coworkers' average quality.

TABLE 1. BASELINE SPECIFICATION ESTIMATES

	Panel A. Horizon in Years (1 to 5 years)				
	1	2	3	4	5
$\hat{\gamma}$	0.141***	0.107***	0.078***	0.064***	0.046***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Within R^2	0.709	0.716	0.728	0.743	0.760
Observations	11,257,738	9,036,667	7,543,899	6,356,797	5,426,863

	Panel B. Horizon in Years (6 to 10 years)				
	6	7	8	9	10
$\hat{\gamma}$	0.048***	0.034***	0.036***	0.034***	0.031***
	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Within R^2	0.775	0.793	0.811	0.828	0.845
Observations	4,655,742	3,962,176	3,350,271	2,776,496	2,237,634

Notes. Estimates of the peer effect parameter from Equation 1's specification. This parameter represents the effect of average peer quality on workers' wages, h years into the future. Peer quality is measured through average individual effects of coworkers. Each column represents a different time horizon h. Controls for age, squared age, tenure, squared tenure, schooling, and squared schooling were included in all specifications. Standard errors clustered at the worker level. Significance levels are represented by stars: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parenthesis.
Source. Quadros de Pessoal, 2002-2021.

Comparing our results to those reported in other studies is not straightforward. As stated before, most of the non-experimental literature on the topic of peer effects in the workplace focuses on specific occupations or firms, meaning that their estimates are not comparable to those obtained in our analysis (which studies peer effects in the broader labour force). Moreover, while the literature tends to concentrate on contemporaneous peer effects, our analysis focuses on dynamic peer effects (i.e., the effect of peers on future wages). As such, and to facilitate comparability, in addition to our baseline results, we estimated the contemporaneous peer effect using our empirical framework (Table A.7). Essentially, this means we estimated Equation 1 for $h = 0$.

The contemporaneous specification yields a coefficient of 0.182, meaning that a one standard deviation increase in current average coworker quality increases contemporaneous wages by 5.2% ($[e^{0.182}-1] \times 0.26$). This coefficient is somewhere between the coefficients obtained by Battisti

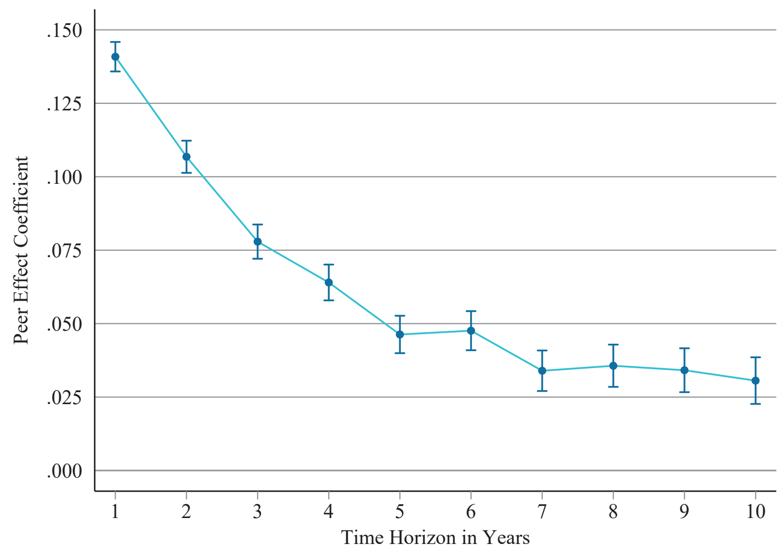


FIGURE 2. PEER EFFECT ON FUTURE WAGES

Notes. The figure graphically represents the estimates from Table 1. The dots correspond to the coefficient estimates across the different time horizons. The vertical lines plot the 95% confidence intervals of each estimate.
Source. Quadros de Pessoal, 2002-2021.

(2017) for the Veneto region (Italy) - 7.8% - and those obtained by Cornelissen, Dustmann, and Schönberg (2017) for the region of Munich (Germany) - 0.3%. Our coefficient is comparable to the coefficient reported by Hong and Lattanzio (2022) - 4.6% - who, like Battisti, exploited data for the Veneto region. In addition to this, our coefficient is naturally very close to the one presented by Portugal et al. (2022) - 5.7% - as both were obtained from an analysis of Quadros de Pessoal.[14] All the above studies, like ours, use the framework developed by Arcidiacono et al. (2012).

VI.B. Disentanglement of Mechanisms

As discussed in Section III.B, the literature contemplates two mechanisms through which peer effects may arise in the workplace. If, on the one hand, a positive peer effect may find its root in the accumulation of human capital resulting from knowledge transmission among peers (knowledge spillovers). On the other, this same positive effect may result from the increase in a worker's effort

14. Even though, they focus their analysis on a different period, 1994 to 2013.

upon observing coworkers' behaviour (peer pressure). In this case, our findings of persistent peer effects across time, give clear support to the notion of knowledge spillovers.

Social interactions between coworkers allow for the implementation of ideas and better practices (that increase productivity) which workers may adopt upon exposure, hence internalizing the peer effect on their own ability. This process of knowledge spillover is thus likely to be persistent over time. As for peer pressure, the impact is not persistent - canonically, a worker cannot be subject to peer pressure from past peers, only from contemporaneous ones. As such, whereas knowledge transmission among peers is likely to still impact wages in subsequent periods (in line with our results), peer pressure, by definition, only materializes contemporaneously.

Note, however, that this does not dismiss the existence of peer pressure. As a matter of fact, the highest estimate is observed for the contemporaneous specification (Table A.7). Nevertheless, as we are not able to disentangle the mechanisms for this particular specification it remains unclear whether contemporaneous peer effects are fully explained by knowledge spillovers or if peer pressure may explain a part of these. To further test the robustness of peer effect persistence we separately analyze *movers* (who changed firms between t and $t + h$) and *stayers* in Section VI.D.

VI.C. Heterogeneous Effects Results

Table 2 reports the estimates for the specification that allows for heterogeneous responsiveness to peers (Equation 3). In this model, the responses to peer quality are allowed to vary according to a worker's own ability-specificity - high or low. Panel A reports the estimates for the response to peer effects by high-ability workers. Panel B reports the same response coefficient but for low-ability workers. Though not having significantly different responses, low-ability workers seem to be more responsive to increases in average peer quality than high-ability workers. But, the difference seems rather small for any conclusion to be drawn from it.[15] Figure 3a plots these results.

15. Note that inference in this Section is limited. We obtained estimates for the heterogeneous effects models by manipulating the fixed effects from the baseline for each time horizon. This implies that the estimated standard errors do not account for the loss in degrees of freedom. A solution to this problem would be the implementation of a bootstrapping procedure. This exercise was not performed due to computational power limitations.

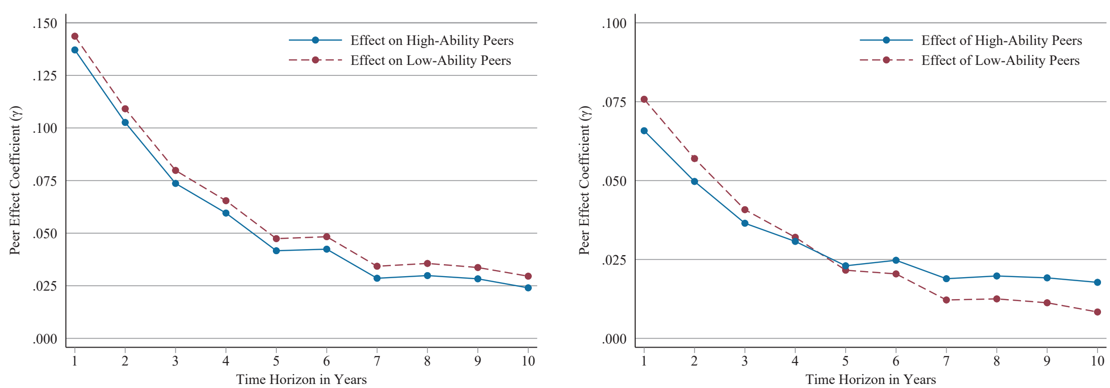
TABLE 2. HETEROGENEOUS RESPONSE TO PEERS ESTIMATES

	Panel A. High-Quality Peers (Top 50%)				
	1	2	3	5	10
$\hat{\gamma}$	0.137	0.103	0.074	0.042	0.024
Observations	4,266,361	3,465,550	2,907,405	2,091,885	875,833

	Panel B. Low-Quality Peers (Bottom 50%)				
	1	2	3	5	10
$\hat{\gamma}$	0.144	0.109	0.080	0.047	0.030
Observations	4,371,152	3,554,300	2,982,416	2,147,298	896,974

Notes. Estimates resulting from implementing Equation 3. These coefficients quantify the heterogeneous response of ability-specific workers to increases in average coworker quality. Panel A reports the response by high-quality workers, Panel B the response by low-ability workers. Workers are divided into high and low-ability subgroups according to each one's individual fixed effect in comparison to their peers' (obtained from the baseline specification). Each column represents a different time horizon h. Controls for age, squared age, tenure, squared tenure, schooling, and squared schooling were included in all specifications. Results for all time horizons can be checked in Table A.5.
Source. Quadros de Pessoal, 2002-2021.

FIGURE 3. HETEROGENEITY IN PEER EFFECTS



(A) RESPONSE TO PEERS (B) PEER INFLUENCE

Notes. The figures plot heterogeneity in peer effects along two dimensions. Figure 3a represents the heterogeneity in response to peers whose estimates were retrieved from Table 2. Figure 3b represents the heterogeneity of peer influence whose estimates were retrieved from Panel A of Table 3.
Source. Quadros de Pessoal, 2002-2021.

TABLE 3. HETEROGENEOUS PEER INFLUENCE ESTIMATES

	Panel A. Full Sample				
	1	2	3	5	10
$\hat{\gamma}^+$	0.066	0.050	0.036	0.023	0.018
$\hat{\gamma}^-$	0.076	0.057	0.041	0.022	0.008
Observations	8,637,892	7,020,071	5,889,972	4,239,302	1,772,839

	Panel B. High-Quality Peers (Top 50%)				
	1	2	3	5	10
$\hat{\gamma}^+$	0.070	0.053	0.039	0.025	0.017
$\hat{\gamma}^-$	0.065	0.049	0.034	0.016	0.006
Observations	4,266,361	3,465,550	2,907,405	2,091,885	875,833

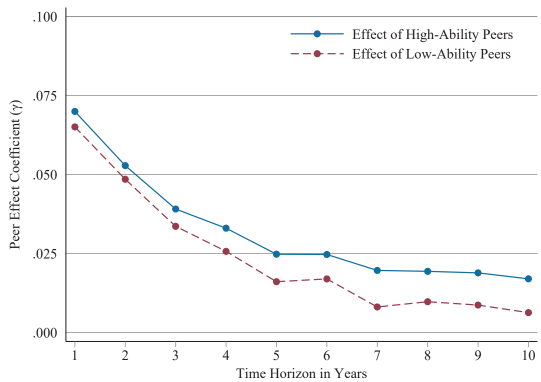
	Panel C. Low-Quality Peers (Bottom 50%)				
	1	2	3	5	10
$\hat{\gamma}^+$	0.061	0.046	0.034	0.021	0.019
$\hat{\gamma}^-$	0.085	0.065	0.047	0.027	0.010
Observations	4,371,152	3,554,300	2,982,416	2,147,298	896,974

Notes. Estimates resulting from implementing Equations 5 and 6. These coefficients quantify the heterogeneous peer influence of each ability-specific peer subgroup. Panel A reports the response to heterogeneous peer influence by the full sample. Panel B reports the response by high-ability workers. Panel C the response by low-ability workers. Each column represents a different time horizon h. Controls for age, squared age, tenure, squared tenure, schooling, and squared schooling were included in all specifications. Results for all time horizons are reported in Tables **??** and A.6.
Source. Quadros de Pessoal, 2002-2021.

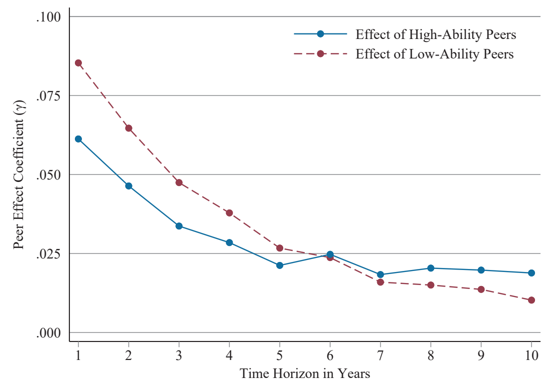
Table 3 presents the estimates for the peer model that allows for heterogeneous peer influence (Equation 5). Here, the baseline model is extended to allow the peer effect to vary according to peers' ability-specificity. Panel A reports the results for the full sample. Panels B and C report the estimates for the specification that also allows ability-specific heterogeneous responses (Equation 6). Panel B presents estimates for the response by high-ability workers to the effect of high (γ^+) and low-ability peers (γ^-). Panel C reports the same responses but by low-ability workers. Figure 3b plots the estimates in Panel A. Figure 4 plots the estimates in Panels B and C.

From Panel A's estimates, we observe a particularly interesting evolution. The coefficient (for

FIGURE 4. ABILITY-SPECIFIC RESPONSES TO HETEROGENEOUS PEER INFLUENCE



(A) BY HIGH-ABILITY WORKERS



(B) BY LOW-ABILITY WORKERS

Notes. The figures plot the different responses of each ability-specific group to heterogeneous peer influence. The estimates were retrieved from Panels A and B of Table 3.
Source. Quadros de Pessoal, 2002-2021.

peer influence) of low-ability peers is higher than that of high-ability ones for the first years post-exposure. Nevertheless, after the fourth year, this is inverted and the coefficient for high-ability peers' influence becomes higher. Put differently, an increase in the contemporaneous peer quality of low-ability peers yields a higher return (in workers' future wages) in the short run. But, in the long run, the return from increasing the quality of high-ability peers seems to yield a higher return.

By closely looking at Figure 3b, we understand that this particularity is explained by the fact that the influence of high-ability peers is remarkably more persistent across time than the influence of low-ability peers. While the peer effect of high-ability peers in the tenth year is roughly one-quarter of the effect after one year, the peer effect of low-ability peers after 10 years is only one-tenth of the effect after one year. Furthermore, even when distinguishing ability-specific responses (Panels B and C of Table 3), we see a similar case regarding the persistence of ability-specific peer effects. For high-ability (low-ability) workers, after 10 years the values for γ^{+} and for γ^{-}, respectively, fall to one-fourth (one-third) and one-tenth of their value after one year. This result may be due to the distinct nature of the knowledge being transmitted by each ability-specific peers.

In addition to this, when considering the ability-specific responses to heterogeneous peer effects, we see yet another interesting phenomenon. Workers seem to be more responsive to peer influence by similar peers. In other words, for high-quality workers, an increase in the quality of high-ability peers yields a higher return than an increase in the quality of low-ability peers. The same applies to low-ability workers, even though the ability-specific heterogeneous persistence leads to a situation in which an increase in the quality of high-ability peers also yields a higher return to low-ability workers after the sixth year.

These different responses may suggest some degree of ability-related segregation within peer groups. High-ability coworkers tend to respond more and influence more fellow high-ability peers. Similarly, low-ability coworkers also tend to respond more and influence more fellow low-ability peers. This is a particularly interesting result, given the already quite granular peer groups we have defined.[16] This may reveal that even when considering small peer groups, workers tend to more intensely interact with and learn from similar peers. A result that is not necessarily new, but that further reinforces the idea of homophily in human relations.

VI.D. Robustness Checks

Different Fixed Effects Specifications. To evaluate the reliability of our baseline estimates, we performed two exercises of robustness checking. The first focuses on allowing alternative levels of heterogeneity. In other words, we control for different fixed effects specifications. Table A.8 reports the results from this exercise. Here, we see that more aggregate fixed effects specifications generate higher peer effects coefficients. This is mostly explained by the less rigid control of sorting phenomena in these alternative specifications. At the same time, the decaying trend in peer effects (as the time horizon increases) is verified across specifications.

Movers and Stayers. We perform a second robustness checking exercise, where we analyze the evolution of the response to peer effects by *movers* and by *stayers*. The goal of such an exercise is to verify whether the results of peer effects persistence are robust to job separations. For this,

16. For reference, the average peer group size in our sample is 4.2. In Cornelissen, Dustmann, and Schönberg (2017) average peer group size is 9.3, in Hong and Lattanzio (2022) 12, and in Portugal et al. (2022) 4.9.

we define as mover any worker that changes firms from t to $t+h$, and as stayer any worker that remained in the same firm during this period. To implement this, we use the framework laid out in Equation 3. Table A.5 presents the estimates of this exercise. We confirm that the persistence of peer effects in future wages is robust to whether a worker remains in the same firm or not.

VII. CONCLUSION

In this study, we analyzed the hypothesis of peer quality influencing workers' future outcomes. To exploit this question, we implemented an empirical strategy that addressed common sources of endogeneity and applied it to a unique longitudinal matched employer-employee dataset. We believe that the obtained estimates are reasonably precise, given our rigid identification strategy.

We found that a ceteris paribus one standard deviation increase in the average quality of contemporaneous coworkers, increases next year's workers' wages by 4.8%. We also found evidence that, although decaying, peer effects are persistent across time. So much so, that we report significant peer effects even 10 years after exposure to better peers. This persistent effect supports the notion that peer effects materialize through knowledge spillovers.

Finally, we also explore the heterogeneity of peer effects. Our results suggest that increasing the average quality of low-ability coworkers yields a higher return (in the form of future wages) in the short run, whereas an increase in the quality of high-ability workers yields a higher return in the long run. This result is explained by the higher persistence of high-ability peers' influence. We also obtained results suggestive of ability-specific segregation within peer groups: high-ability (low-ability) workers tend to, more intensely, influence and respond to fellow high-ability (low-ability) peers. A result that reinforces the idea of homophily in social interactions and human relations.

References

Abowd, John M., Robert H. Creecy, and Francis Kramarz. 2002. *Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data.* Technical Paper. US Census Bureau.

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica* 67 (2): 251–333.

Angrist, Joshua D. 2014. "The perils of peer effects." *Labour Economics* 30:98–108.

Arcidiacono, Peter, Gigi Foster, Natalie Goodpaster, and Josh Kinsler. 2012. "Estimating Spillovers Using Panel Data, with an Application to the Classroom." *Quantitative Economics* 3 (3): 421–470.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2010. "Social Incentives in the Workplace." *The Review of Economic Studies* 77 (2): 417–458.

Battisti, Michele. 2017. "High Wage Workers and High Wage Peers." *Labour Economics* 46:47–63.

Bayer, Patrick, Randi Hjalmarsson, and David Pozen. 2009. "Building Criminal Capital behind Bars: Peer Effects in Juvenile Corrections." *The Quarterly Journal of Economics* 124 (1): 105–147.

Borjas, George J., and Kirk B. Doran. 2015. "Which Peers Matter? The Relative Impacts of Collaborators, Colleagues, and Competitors." *The Review of Economics and Statistics* 97 (5): 1104–1117.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin. 2009. "Identification of Peer Effects through Social Networks." *Journal of Econometrics* 150 (1): 41–55.

Chetty, Raj, and Nathaniel Hendren. 2018. "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects." *The Quarterly Journal of Economics* 133 (3): 1107–1162.

Cornelissen, Thomas, Christian Dustmann, and Uta Schönberg. 2017. "Peer Effects in the Workplace." *American Economic Review* 107 (2): 425–456.

Davidson, Russell, and James G. MacKinnon. 2006. "Bootstrap Methods in Econometrics." In *Palgrave Handbook of Econometrics, Volume 1: Econometric Theory,* edited by Terence C. Mills and Kerry Patterson, 812–838. Basingstoke: Palgrave Macmillan.

FALK, ARMIN, and ANDREA ICHINO. 2006. "Clean Evidence on Peer Effects." *Journal of Labor Economics* 24 (1): 39–57.

GLAESER, EDWARD L., BRUCE SACERDOTE, and JOSÉ A. SCHEINKMAN. 1996. "Crime and Social Interactions." *The Quarterly Journal of Economics* 111 (2): 507–548.

GURYAN, JONATHAN, KORY KROFT, and MATTHEW J. NOTOWIDIGDO. 2009. "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments." *American Economic Journal: Applied Economics* 1 (4): 34–68.

HONG, LONG, and SALVATORE LATTANZIO. 2022. "The Peer Effect on Future Wages in the Workplace." SSRN Working Paper.

JACKSON, C. KIRABO, and ELIAS BRUEGMANN. 2009. "Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers." *American Economic Journal: Applied Economics* 1 (4): 85–108.

JAROSCH, GREGOR, EZRA OBERFIELD, and ESTEBAN ROSSI-HANSBERG. 2021. "Learning From Coworkers." *Econometrica* 89 (2): 647–676.

LUCAS, ROBERT E. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (1): 3–42.

MANSKI, CHARLES F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies* 60 (3): 531–542.

MARSHALL, ALFRED. 1890. *Principles of Economics.* London: Macmillan.

MAS, ALEXANDRE, and ENRICO MORETTI. 2009. "Peers at Work." *American Economic Review* 99 (1): 112–145.

OREOPOULOS, PHILIP. 2003. "The Long-Run Consequences of Living in a Poor Neighborhood." *The Quarterly Journal of Economics* 118 (4): 1533–1575.

PORTUGAL, PEDRO, HUGO REIS, PAULO GUIMARÃES, and ANA RUTE CARDOSO. 2022. "Human Capital Spillovers and Returns to Education." Banco de Portugal Working Paper.

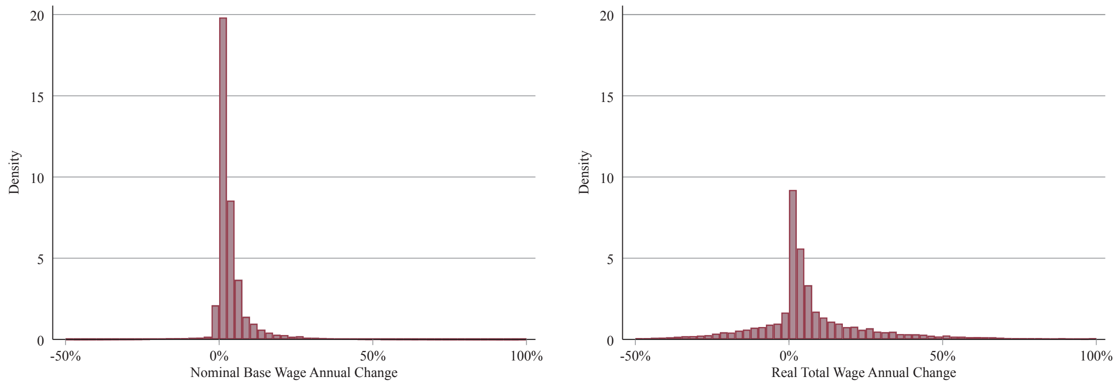
ROMER, PAUL M. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98 (5): S71–S102.

SACERDOTE, BRUCE. 2011. "Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" In *Handbook of the Economics of Education,* edited by Erik Hanushek, Stephen Machin, and Ludger Woessmann, 249–277. Amsterdam: Elsevier.

WALDINGER, FABIAN. 2012. "Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany." *The Review of Economic Studies* 79 (2): 838–861.

Appendix

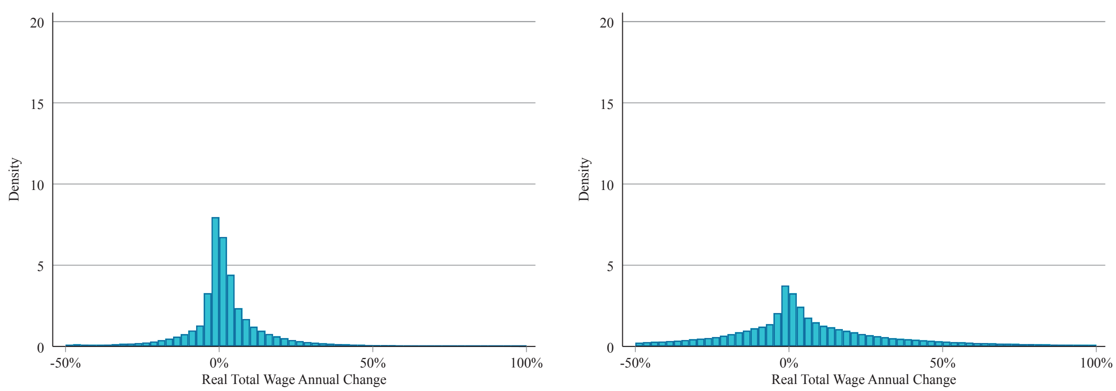
Figure A.1. Histogram of the Annual Change of Nominal Base Wage, Variations Across the Mobility Dimension



(A) Stayers (B) Movers

Notes. The figures plot the histograms of yearly growth rates of nominal base wages. Each plot represents a different mobility status that is identified in the plot's caption. The rigidity of nominal base wages is much more apparent for workers who remain in the same firm. Movers observe a certain wage flexibility as their wages result from a new labour contract. The figure is trimmed at the -50% and 100% thresholds.
Source. Quadros de Pessoal, 2002-2021.

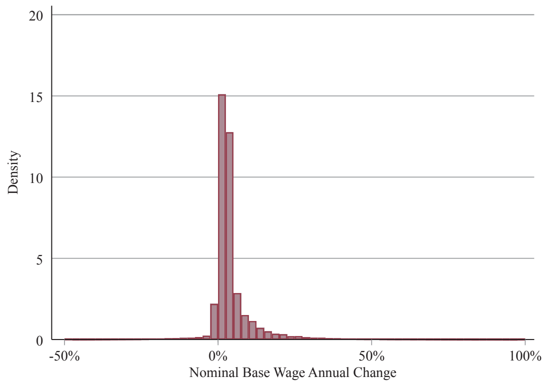
Figure A.2. Histogram of the Annual Change of Real Total Wage, Variations Across the Mobility Dimension



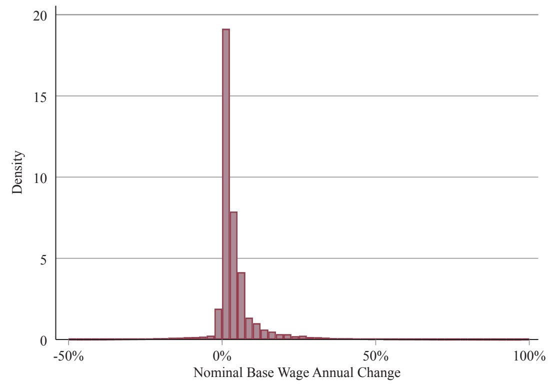
(A) Stayers (B) Movers

Notes. The figures plot the histograms of yearly growth rates of real total wages. Each plot represents a different mobility status that is identified in the plot's caption. Under this measure of labour remuneration, wage rigidity is significantly attenuated for both situations of job mobility. The figure is trimmed at the -50% and 100% thresholds.
Source. Quadros de Pessoal, 2002-2021.

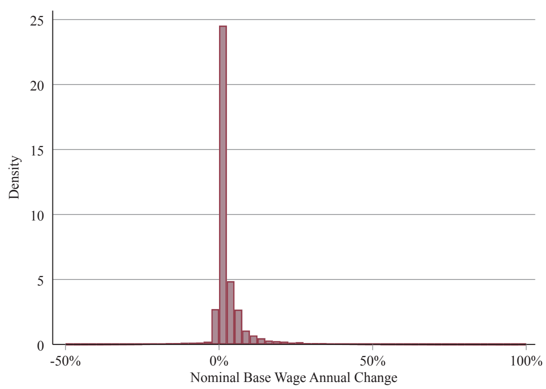
FIGURE A.3. HISTOGRAM OF THE ANNUAL CHANGE OF NOMINAL BASE WAGE, VARIATIONS ACROSS THE TIME DIMENSION



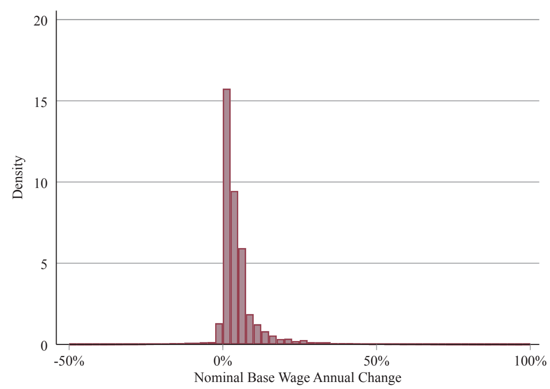
(A) 2002-2006



(B) 2007-2011



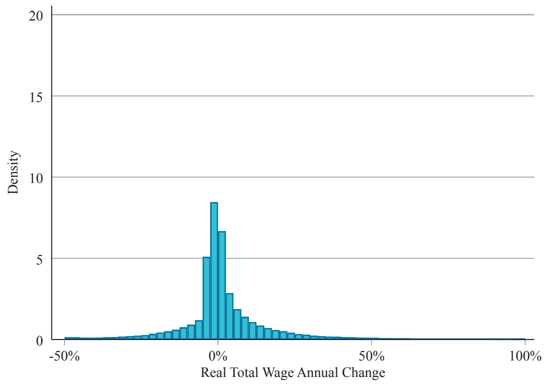
(C) 2012-2016



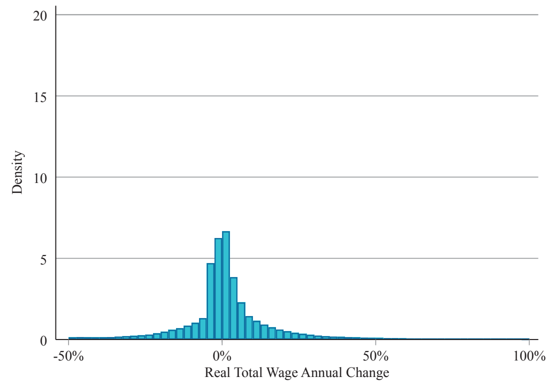
(D) 2017-2021

Notes. The figures plot the histograms of yearly growth rates of nominal base wages. Each plot represents a different time frame that is identified in the plot's caption. The remarkable nominal base wage rigidity persists across the four different time frames. Note that the y-axis scale of the third plot is different than the scale of the other three. The figure is trimmed at the -50% and 100% thresholds.
Source. Quadros de Pessoal, 2002-2021.

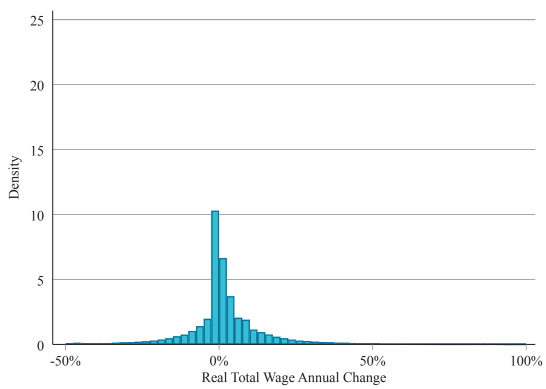
FIGURE A.4. HISTOGRAM OF THE ANNUAL CHANGE OF REAL TOTAL WAGE, VARIATIONS ACROSS THE TIME DIMENSION



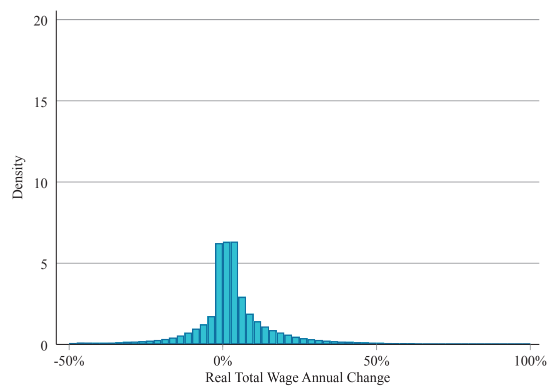
(A) 2002-2006



(B) 2007-2011



(C) 2012-2016



(D) 2017-2021

Notes. The figures plot the histograms of yearly growth rates of real total wages. Each plot represents a different time frame that is identified in the plot's caption. Under this measure of labour remuneration, the previously observed rigidity of nominal base wages is significantly attenuated. Note that the y-axis scale of the third plot is different than the scale of the other three. The figure is trimmed at the -50% and 100% thresholds.
Source. Quadros de Pessoal, 2002-2021.

TABLE A.1. ANNUAL CHANGE OF DIFFERENT WAGE MEASURES (IN %)

	Mean	Std. Dev.	25th Perc.	Median	75th Perc.
Nominal Base Wage	4.03	12.29	0.00	2.00	4.72
Stayers	*3.71*	*10.40*	*0.00*	*1.95*	*4.62*
Movers	*9.17*	*28.77*	*0.00*	*3.31*	*13.16*
2002-2006	*4.60*	*12.96*	*0.00*	*2.52*	*4.32*
2007-2011	*4.17*	*13.02*	*0.00*	*2.10*	*4.94*
2012-2016	*2.43*	*10.92*	*0.00*	*0.00*	*2.44*
2016-2021	*5.05*	*12.17*	*0.00*	*3.39*	*5.69*
Real Total Wage	*4.58*	*26.29*	*-2.43*	*0.69*	*6.75*
Stayers	*4.17*	*24.34*	*-2.33*	*0.65*	*6.28*
Movers	*11.34*	*47.54*	*-9.24*	*2.25*	*21.44*
2002-2006	*4.19*	*27.31*	*-2.61*	*-0.03*	*5.72*
2007-2011	*4.14*	*27.84*	*-3.60*	*0.60*	*6.46*
2012-2016	*3.97*	*24.64*	*-2.58*	*0.26*	*6.10*
2016-2021	*5.89*	*25.47*	*-1.36*	*2.39*	*8.12*
		Observations	12,431,437		
		Stayers	11,717,309		
		Movers	714,128		
		2002-2006	2,320,356		
		2007-2011	3,397,897		
		2012-2016	3,397,897		
		2016-2021	3,403,303		

Notes. The table reports summary statistics on the yearly variation in the two wage measures - nominal base wages and real total wages - we compared in Figure 1. The table reports statistics across the different dimensions analyzed in Figures A.1 to A.4. The values are reported in percentages, except for the different numbers of observations. Note that the 2002-2006 timeframe has a significantly lower number of observations than other timeframes as we have no observations on yearly change for 2002 (as there was no data collected by QP in 2001).
Source. Quadros de Pessoal, 2002-2021.

TABLE A.2. LARGEST CONNECTED SET FOR DIFFERENT TIME HORIZONS

	Full Sample	Connected Observations	Non-Connected Observations	Coverage Ratio
0-Year Time Horizon	20,065,301	18,593,626	1,471,675	92.67%
1-Year Time Horizon	13,210,891	11,257,738	1,953,153	85.22%
2-Year Time Horizon	10,918,670	9,036,667	1,882,003	82.76%
3-Year Time Horizon	9,283,098	7,543,899	1,739,199	81.26%
4-Year Time Horizon	7,947,166	6,356,797	1,590,369	79.99%
5-Year Time Horizon	6,879,737	5,426,863	1,452,874	78.88%
6-Year Time Horizon	5,986,420	4,655,742	1,330,678	77.77%
7-Year Time Horizon	5,192,229	3,962,176	1,230,053	76.31%
8-Year Time Horizon	4,503,669	3,350,271	1,153,398	74.39%
9-Year Time Horizon	3,879,208	2,776,496	1,102,712	71.57%
10-Year Time Horizon	3,312,532	2,237,634	1,074,898	67.55%

Notes. The table reports the largest connected set of workers and peer groups in Portugal for each of the time horizons analyzed in our regressions (1 year to 10 years) and for the simple contemporaneous time horizon. Sample selection details can be found in Section V.
Source. Quadros de Pessoal, 2002-2021.

TABLE A.3. SUMMARY STATISTICS

	Mean	Std. Dev.	Min	Max
Real Base Wage (in 2021 Euros)	902.05	581.92	366.01	6,497.73
2002-2006	*868.50*	*579.97*	*366.01*	*6,374.25*
2007-2011	*906.62*	*607.47*	*374.33*	*6,497.73*
2012-2016	*890.76*	*579.87*	*409.94*	*6,091.95*
2016-2021	*938.93*	*556.63*	*458.49*	*5,910.00*
Real Total Wage (in 2021 Euros)	1,188.68	788.50	366.01	11,304.81
2002-2006	*1,128.25*	*779.71*	*366.01*	*11,304.81*
2007-2011	*1,187.95*	*814.27*	*374.41*	*10,835.60*
2012-2016	*1,180.57*	*789.23*	*409.94*	*10,906.93*
2016-2021	*1,252.67*	*764.37*	*458.49*	*10,587.11*
Real Hourly Wage (in 2021 Euros)	6.99	4.77	2.04	37.19
2002-2006	*6.69*	*4.76*	*2.04*	*35.10*
2007-2011	*6.99*	*4.93*	*2.13*	*37.19*
2012-2016	*6.95*	*4.79*	*2.37*	*35.42*
2016-2021	*7.32*	*4.58*	*2.65*	*34.20*
Monthly Hours Worked	171.45	10.87	120	421
Age (in Years)	39.16	10.73	18	64
Tenure (in Years)	8.37	8.80	0	50
Years of Schooling	9.24	3.92	0	23
Peer Group Size	4.18	3.98	2	33
Firm Size	11.13	67.88	2	13,251

Worker-Year Observations	18,593,626
2002-2006	4,433,698
2007-2011	4,840,092
2012-2016	4,506,852
2016-2021	4,812,984
Peer Group-Year Observations	4,446,157
Firm-Year Observations	1,669,946
Worker Observations	3,518,197
Peer Group Observations	1,753,795
Firm Observations	287,133

Notes. The table reports descriptive statistics for real wage variables and other relevant covariates. Statistics for wage variables are also reported by 5-year intervals. These summary statistics are based on the largest connected set of workers and peer groups in Portugal. Sample selection details can be found in Section V.
Source. Quadros de Pessoal, 2002-2021.

TABLE A.4. SUMMARY STATISTICS ON WORKER AND PEER FIXED EFFECTS

	Mean	Std. Dev.	Min	Max	Observations
Worker Fixed Effects					
0-Year Time Horizon	0.00	0.29	-3.79	3.92	18,593,626
1-Year Time Horizon	0.00	0.35	-3.17	3.62	11,257,738
2-Year Time Horizon	0.00	0.39	-3.33	3.39	9,036,667
3-Year Time Horizon	0.00	0.44	-3.84	3.49	7,543,899
4-Year Time Horizon	0.00	0.47	-3.27	3.92	6,356,797
5-Year Time Horizon	0.00	0.49	-3.45	4.61	5,426,863
6-Year Time Horizon	0.00	0.51	-3.39	3.94	4,655,742
7-Year Time Horizon	0.00	0.52	-3.79	3.70	3,962,176
8-Year Time Horizon	0.00	0.52	-3.11	3.40	3,350,271
9-Year Time Horizon	0.00	0.52	-4.03	3.85	2,776,496
10-Year Time Horizon	0.00	0.52	-3.42	4.25	2,237,634
Average Coworker Fixed Effects					
0-Year Time Horizon	0.00	0.26	-3.79	3.92	18,593,626
1-Year Time Horizon	0.00	0.32	-3.17	3.54	11,257,738
2-Year Time Horizon	0.00	0.37	-3.33	3.39	9,036,667
3-Year Time Horizon	0.00	0.41	-3.84	3.49	7,543,899
4-Year Time Horizon	0.00	0.44	-3.27	3.92	6,356,797
5-Year Time Horizon	0.00	0.47	-3.45	4.61	5,426,863
6-Year Time Horizon	0.00	0.48	-3.39	3.94	4,655,742
7-Year Time Horizon	0.00	0.49	-3.79	3.70	3,962,176
8-Year Time Horizon	0.00	0.49	-3.02	3.40	3,350,271
9-Year Time Horizon	0.00	0.49	-4.03	3.85	2,776,496
10-Year Time Horizon	0.00	0.49	-3.37	4.25	2,237,634

Notes. The table reports descriptive statistics on workers' individual fixed effects as well as on average coworkers' fixed effects. These summary statistics refer to the fixed effects estimates obtained from the baseline specification for each time horizon. For further details on the specification and estimation of these fixed effects check Section IV.
Source. Quadros de Pessoal, 2002-2021.

TABLE A.5. HETEROGENEOUS RESPONSE TO PEERS SPECIFICATIONS

Time Horizon		Baseline Full Sample (1)	Ability-Specific H-Ability (2)	Ability-Specific L-Ability (3)	Movers / Stayers Movers (4)	Movers / Stayers Stayers (5)
$h = 1$	$\hat{\gamma}$	0.141	0.137	0.144	0.108	0.142
		11,257,738	*4,266,363*	*4,371,152*	*590,875*	*10,666,863*
$h = 2$	$\hat{\gamma}$	0.107	0.103	0.109	0.084	0.109
		9,036,667	*3,465,550*	*3,554,300*	*959,307*	*8,077,360*
$h = 3$	$\hat{\gamma}$	0.078	0.074	0.080	0.061	0.080
		7,543,899	*2,907,405*	*2,982,416*	*1,159,813*	*6,384,086*
$h = 4$	$\hat{\gamma}$	0.064	0.060	0.065	0.051	0.066
		6,356,797	*2,451,935*	*2,516,094*	*1,239,703*	*5,117,094*
$h = 5$	$\hat{\gamma}$	0.046	0.042	0.047	0.035	0.048
		5,426,863	*2,091,885*	*2,147,298*	*1,259,785*	*4,167,078*
$h = 6$	$\hat{\gamma}$	0.048	0.042	0.048	0.038	0.048
		4,655,742	*1,795,121*	*1,842,982*	*1,239,619*	*3,416,123*
$h = 7$	$\hat{\gamma}$	0.034	0.029	0.034	0.026	0.034
		3,962,176	*1,528,424*	*1,568,431*	*1,178,453*	*2,783,723*
$h = 8$	$\hat{\gamma}$	0.036	0.030	0.036	0.028	0.035
		3,350,271	*1,293,661*	*1,328,379*	*1,093,982*	*2,256,289*
$h = 9$	$\hat{\gamma}$	0.034	0.028	0.034	0.027	0.033
		2,776,496	*1,077,486*	*1,105,048*	*983,329*	*1,793,167*
$h = 10$	$\hat{\gamma}$	0.034	0.024	0.030	0.024	0.029
		2,237,634	*875,833*	*896,974*	*856,791*	*1,380,843*

Notes. Estimates of the heterogeneous response to peers specifications. The rows represent the different time horizons h. Each column represents the different response-to-peers subgroups. In this Table, we include the responses of ability-specific subgroups as well as the responses of movers and stayers. The division into ability-specific subgroups is detailed in Section IV. *H-Ability* refers to the responses by high-ability individuals, *L-ability* the response by low-ability ones. Controls for age, squared age, tenure, squared tenure, schooling, and squared schooling were included in all specifications. Number of observations in italics.
Source. Quadros de Pessoal, 2002-2021.

TABLE A.6. HETEROGENEOUS PEER INFLUENCE ESTIMATES

Time Horizon		Full Sample (1)	Halves: H-Ability (2)	Halves: L-Ability (3)
$h = 1$	$\hat{\gamma}^+$	0.066	0.070	0.061
	$\hat{\gamma}^-$	0.076	0.065	0.085
		8,637,892	*4,266,361*	*4,371,152*
$h = 2$	$\hat{\gamma}^+$	0.050	0.053	0.046
	$\hat{\gamma}^-$	0.057	0.049	0.065
		7,020,071	*3,465,550*	*3,554,300*
$h = 3$	$\hat{\gamma}^+$	0.036	0.039	0.034
	$\hat{\gamma}^-$	0.041	0.034	0.047
		5,889,972	*2,907,405*	*2,982,416*
$h = 4$	$\hat{\gamma}^+$	0.031	0.033	0.028
	$\hat{\gamma}^-$	0.032	0.026	0.038
		4,968,161	*2,451,935*	*2,516,094*
$h = 5$	$\hat{\gamma}^+$	0.023	0.025	0.021
	$\hat{\gamma}^-$	0.022	0.016	0.027
		4,239,302	*2,091,885*	*2,147,298*
$h = 6$	$\hat{\gamma}^+$	0.025	0.025	0.025
	$\hat{\gamma}^-$	0.020	0.017	0.024
		3,638,184	*1,795,121*	*1,842,982*
$h = 7$	$\hat{\gamma}^+$	0.019	0.020	0.018
	$\hat{\gamma}^-$	0.012	0.008	0.016
		3,096,923	*1,528,424*	*1,568,431*
$h = 8$	$\hat{\gamma}^+$	0.020	0.019	0.020
	$\hat{\gamma}^-$	0.013	0.010	0.015
		2,622,100	*1,293,661*	*1,328,379*
$h = 9$	$\hat{\gamma}^+$	0.019	0.019	0.020
	$\hat{\gamma}^-$	0.011	0.009	0.014
		2,182,593	*1,077,486*	*1,105,048*
$h = 10$	$\hat{\gamma}^+$	0.018	0.017	0.019
	$\hat{\gamma}^-$	0.008	0.006	0.010
		1,772,839	*875,833*	*896,974*

Notes. Estimates of the heterogeneous peer influence specifications. The rows represent the different time horizons h. For each time horizon, we report the high-ability peers' influence, γ^+, and the low-ability peers' influence γ^-. Each column represents the different response-to-peers subgroups. The division into ability-specific subgroups is detailed in Section IV. *H-Ability* refers to the responses by high-ability individuals, *L-ability* the response by low-ability ones. Controls for age, squared age, tenure, squared tenure, schooling, and squared schooling were included in all specifications. Number of observations in italics.
Source. Quadros de Pessoal, 2002-2021.

TABLE A.7. CONTEMPORANEOUS PEER EFFECTS

	Horizon in Years
	0
$\hat{\gamma}$	0.182^{***}
	(0.002)
Within R^2	0.760
Observations	18,593,626
Worker FE	Yes
Establishment $\times$ Job Title $\times$ Year FE	Yes

Notes. Estimate of the contemporaneous peer effect parameter. The value was obtained by implementing Equation 1 for $h = 0$. Controls for age, squared age, tenure, squared tenure, schooling, and squared schooling were included. Standard errors clustered at the worker level. Significance levels are represented by stars: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parenthesis.
Source. Quadros de Pessoal, 2002-2021.

TABLE A.8. PEERS EFFECTS ESTIMATES ACROSS DIFFERENT SPECIFICATIONS

		Baseline	Alternative Specifications		
		(1)	(2)	(3)	(4)
$h = 1$	$\hat{\gamma}$	0.141***	0.187***	0.317***	0.637***
		(0.003)	(0.001)	(0.001)	(0.001)
		11,257,738	*11,257,738*	*11,257,738*	*11,257,738*
$h = 2$	$\hat{\gamma}$	0.107***	0.134***	0.255***	0.458***
		(0.003)	(0.001)	(0.001)	(0.001)
		9,036,667	*9,036,667*	*9,036,667*	*9,036,667*
$h = 3$	$\hat{\gamma}$	0.078***	0.090***	0.205***	0.313***
		(0.003)	(0.001)	(0.002)	(0.001)
		7,543,899	*7,543,899*	*7,543,899*	*7,543,899*
$h = 4$	$\hat{\gamma}$	0.064***	0.066***	0.165***	0.214***
		(0.003)	(0.001)	(0.002)	(0.001)
		6,391,612	*6,391,612*	*6,391,612*	*6,391,612*
$h = 5$	$\hat{\gamma}$	0.046***	0.050***	0.139***	0.151***
		(0.003)	(0.001)	(0.002)	(0.001)
		5,458,415	*5,458,415*	*5,458,415*	*5,458,415*
Worker FE		Yes	Yes	Yes	Yes
Year FE				Yes	
Estab. × Job T. FE				Yes	
Estab. × Year FE			Yes		
Job T. × Year FE			Yes		
Estab. × Job T. × Year FE		Yes			

Notes. Estimates of the peer effect parameter under different fixed effects specifications. The different time horizons h are represented in the first five rows of the table, after headings. Each column represents a different fixed effects specification, whose particular details are provided in the bottom rows. *Estab.* refers to Establishment, *Job T.* to Job Title, and *FE* to Fixed Effects. The leftmost specification is our baseline specification, representing the most restrictive specification, the second leftmost the second most restrictive, and so on. The rightmost specification is the most relaxed one. The preferred specification, (1), yields the lowest estimates as it controls most extensively for the different sources of sorting. Conversely, the most relaxed specification, (4), yields the highest estimates as it does not control for any source of sourcing besides homophily, through worker fixed effects. Due to computational power limitations, the estimates for the second specification were not estimated at the same precision level as the other specifications'. Controls for age, squared age, tenure, squared tenure, schooling, and squared schooling were included in all specifications. Significance levels are represented by stars: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parenthesis. Number of observations in italics.
Source. Quadros de Pessoal, 2002-2021.

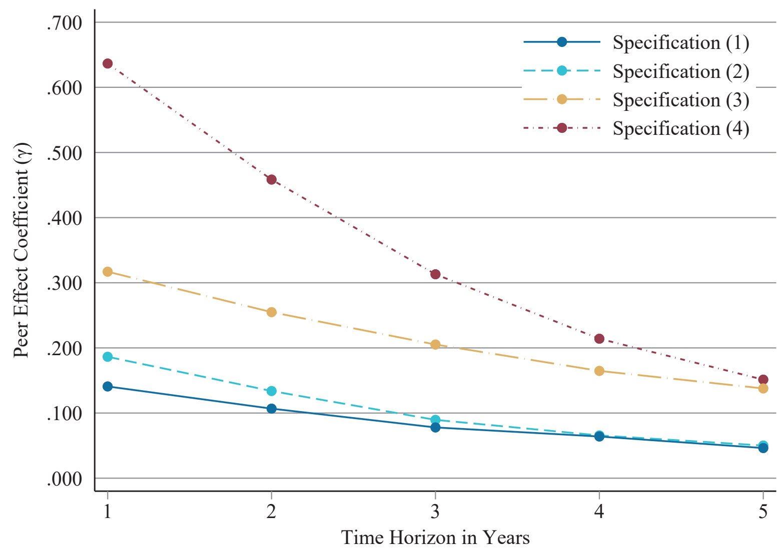


FIGURE A.5. PEER EFFECTS COEFFICIENTS ACROSS DIFFERENT SPECIFICATIONS

Notes. The figure graphically represents the estimates for the different fixed effects specifications reported in Table A.8. Controlling for different sources of sorting significantly reduces estimates. The decaying profile of peer effects is robust across specifications.
Source. Quadros de Pessoal, 2002-2021.

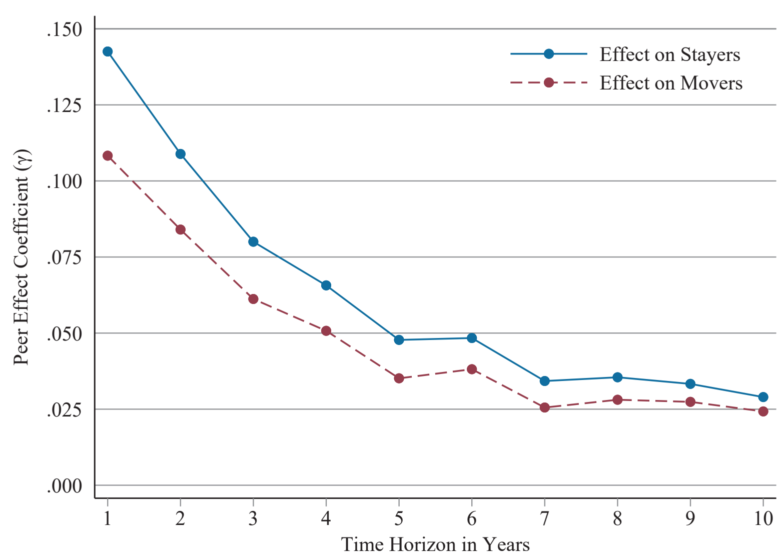


FIGURE A.6. RESPONSE TO PEERS BY MOVERS AND STAYERS

Notes. The figure graphically represents the estimates for the job mobility-specific responses to peers by movers and stayers, reported in Table A.5. Despite the small difference between the coefficients of the two groups, both of them observe significant persistence of peer effects in future wages.
Source. Quadros de Pessoal, 2002-2021.