Mind the Gap: How Coworkers Shape Team Composition (Working paper - Internal version) *

Jaime Montana
Universidade Católica Portuguesa

Miguel Salema
Universidade Católica Portuguesa

Joana Silva
Universidade Católica Portuguesa, World Bank and CEPR

November 23, 2024

Abstract

This paper examines how the effect of the '*peer quality gap*'–the difference between a worker's quality and that of their peers–affects the composition of teams. We leverage matched employer-employee data from Portugal to extract constant worker ability as a proxy for quality and develop a new econometric method to estimate gap effects asymmetrically. With it, we find that among high-skill employees, better workers are more negatively affected by the peer quality gap than worse workers benefit from it. This provides the first mechanism through which the quality gap influences team composition: on the labor demand side, firms benefit from low dispersion in quality due to asymmetric gap effects. Secondly, on the labor supply side, we find that the quality gap incentivizes better employees, who are being negatively affected by worse peers, to move to better teams. The findings emphasize the potential for firms to enhance productivity and minimize turnover by strategically managing team quality composition.

Keywords: Peer effects, Worker quality, Wage determinants, Mobility, Sorting.
JEL Classification: J31, J62, M51

*jdoncel@ucp.pt, misalema@ucp.pt, jsilva@ucp.pt. We are grateful to PEOPLE, and participants in several seminars for helpful comments. We thank the Portuguese Institute of Statistics (INE) for data access. We thank Fundação para a Ciência e a Tecnologia (FCT) for financial support. We declare that we have no relevant or material financial interest that relate to the research described in this paper. We remain responsible for any errors. The views expressed in this paper are those of the authors and should not be attributed to the institutions they are affiliated with.

1 Introduction

Teamwork has revolutionized organizational structures, becoming the backbone of operations in many firms (Lazear and Shaw, 2007). Jobs today rely heavily on collaboration, interaction, and synchronization, underscoring how coworkers significantly shape labor market outcomes. Worker pay and firm productivity, two faces of the same coin, both reflect the profound influence of team dynamics. For workers, having high-quality peers has been linked to better pay (Cornelissen et al., 2017; Cardoso et al., 2022; Messina et al., 2023). For firms, assembling the right team drives measurable productivity gains (Mas and Moretti, 2009; Falk and Ichino, 2006). But what happens when a worker's skills differ from those of their peers? This question–captured in the idea that *'coworkers can make-or-break the job'*–opens a window into the dynamics of job allocation and wage disparities across the labor market.

To understand these dynamics, in this paper, we analyze match quality between a worker and their peers–the '*peer quality gap*'–from both the supply and demand sides of the labor market. Using matched employer-employee data covering 4 million workers across 2.2 million teams in Portugal's private sector, our study provides a uniquely granular perspective. This data allows us to *(i)* define teams based on physical proximity, task similarity, and shared skills, ensuring robust peer group definitions; *(ii)* examine how peer effects differ across worker types, shaping individual and firm-level outcomes; and *(iii)* leverage the panel structure of the data to extract constant worker ability as a proxy for quality .

From the supply side, workers whose quality significantly exceeds their peers may face wage penalties, increasing their likelihood of switching jobs–consistent with an on-the-job search framework (Postel–Vinay and Robin, 2002). These mismatches disproportionately affect groups sensitive to peer effects, underscoring the importance of understanding how peer dynamics vary. Meanwhile, from the demand side, firms aim to maximize profits by strategically allocating workers who are "better" or "worse" than the current team average. Better workers often transition to better jobs, while worse workers are more likely to face worse job opportunities or unemployment. This interplay between worker mobility and peer quality gaps highlights a novel determinant of worker allocation. We show empirically that these dynamics are indeed at play, filling an important gap in the literature with heterogeneous effects on both observable and unobservable worker characteristics.

We examine how the peer quality gap affects wages and ultimately mobility. A key innovation of our study is the use of a constant worker component as a proxy for quality, setting it apart from previous literature. This assumption simplifies the analysis and aligns with theoretical models of peer effects, which often emphasize unobservable intrinsic abilities as key drivers of workplace dynamics. To estimate both individual and average peer quality, we adapt existing empirical approaches to peer effects in the workplace (Cornelissen et al., 2017; Battisti, 2017; Cardoso et al., 2022; Hong and Lattanzio, 2022; Messina et al., 2023). We then calculate the peer quality gap by subtracting individual quality from average peer quality, bringing the model closer to theoretical frameworks of peer pressure (Kandel and Lazear, 1992; Mas and Moretti, 2009; Cornelissen et al., 2017). This novel empirical application reveals a positive relationship between the peer quality gap and wages, with significant implications for worker behavior and firm decision-making.

We conduct our analysis in four steps. First, we benchmark the peer quality gap against existing literature by comparing its predictions to prior studies on peer effects and earnings (Cornelissen et al., 2017; Cardoso et al., 2022; Messina et al., 2023; Hong and Lattanzio, 2022). Our preferred specification accounts for individual worker quality, the interaction of collective agreement with occupation, firm, and year-fixed effects. We find that a one standard deviation increase in the peer quality gap leads to a 12.8% increase in current labor earnings. Moreover, this effect, while persistent, diminishes over time, shrinking to approximately 3.3% after five years. These findings underscore the significant role of peer quality gaps in shaping both current and future wages. Furthermore, the dissipation of the effect over time suggests that knowledge acquisition is unlikely to be the primary mechanism. Instead, this pattern highlights the potential for dissipation to serve as a diagnostic tool for identifying the underlying channels driving peer effects.

We next investigate how peer quality gap effects vary across different populations. Studying heterogeneous effects in large networks poses significant methodological challenges, which we address by introducing a feasible approach for testing these effects. In network theory, similar entities tend to group together (homophily), which can lead to shared behaviors and similar effects within these groups. By grouping workers based on common characteristics, such as demographics, team roles, risk of separation, and mobility status, we identify heterogeneous peer effects between groups.

Females are generally less sensitive to the peer quality gap than males across all age groups, although this gap narrows with age. Younger workers exhibit greater sensitivity to peer quality differences. While the effects for geography and skills are relatively homogeneous, we observe substantial variation in other dimensions. Supervisors are less affected by the peer quality gap, whereas workers at higher risk of separation experience larger effects compared to established workers. Similarly, hires from unemployment are more sensitive to peer quality gaps than poached or incumbent workers. These findings suggest that the peer quality gap operates primarily through peer pressure, with responses varying based on individual vulnerability and team dynamics.

These findings align with broader patterns of wage growth and gender dynamics. Workers, on average, increase their wages by 66% over their careers, with mobility-related factors accounting for a third of this progression. Remarkably, changes in coworkers contribute to two-fifths of this growth– more than double the effect attributed to firm mobility– underscoring the critical role of peer quality in shaping wage trajectories. Moreover, peer quality explains 31% of the gender pay growth gap, making it the largest contributor to this disparity. Collectively, these results suggest that peer quality gaps operate through mechanisms such as peer pressure and individual vulnerability, highlighting their far-reaching implications for wage growth and gender inequality.

In a third step, we examine how the peer quality gap influences team composition. A key question we address is whether 'bad apples' can spoil 'good apples' in team settings. Using a novel approach, we calculate the difference between the effects of better and worse workers across skill levels. Our findings reveal that peer quality gap effects are almost symmetric for low- and medium-skill workers but significantly asymmetric for high-skill workers. For low-skill workers, one worker's loss translates to another's gain, but for high-skill workers, the loss is larger than the gain.

Consider a firm with four high-skill workers split into two teams: one with two high performers (1 standard deviation above the median) and one with two low performers (1 standard deviation below). If one worker from each team switches, creating mixed teams, average wages drop by 1.6% solely due to this reallocation. For low-skill workers, this effect is over three times smaller. These results suggest important implications for team formation: low and medium-skill teams tend to exhibit greater quality dispersion, while high-skill teams are more likely to be homogeneous. This finding aligns with positive assortativity observed in prior empirical studies, where high-performing teams often consist of high-performing workers (Abowd et al., 1999; Battisti, 2017).

Lastly, we explore how the peer quality gap influences worker mobility and sorting in the labor market. By focusing on firms with at least one entry or exit, we calculate the conditional effects of variables such as age, tenure, firm pay premium, residualized wage, and the peer quality gap on the likelihood of worker transitions. Using a Generalized Additive Model (GAM), we analyze how the probability of different types of mobility varies across the distribution of each variable while accounting for the effects of others.

Our findings reveal that the peer quality gap is a significant determinant of mobility. Across all skill levels, workers are more likely to move to better jobs as the peer quality gap increases, with the effect being strongest in the top quartile. Conversely, the likelihood of transitioning to a worse job increases in the bottom quartile of the distribution. For high-skill workers in the bottom quintiles, we also observe a higher likelihood of exiting formal employment altogether, a pattern that may have more pronounced effects on the sorting of high-skill workers. This sorting mechanism suggests that relatively better workers tend to seek better opportunities, while relatively worse workers are more likely to move down or out. These results demonstrate that the peer quality gap not only affects wages but also shapes mobility patterns, reinforcing worker sorting and improving labor market efficiency.

Related Literature Our contribution to the literature is threefold, bridging two key areas of labor economics: peer effects in the workplace and worker mobility. The idea that social connections influence worker outcomes is well established. While sociology has long recognized the role of social ties (Granovetter, 1973, 1983; Montgomery, 1992), economics has more recently quantified their importance in empirical studies. Social ties are known to shape the job-worker matching process (Cahuc and Fontaine, 2009). and influence wages and career trajectories (Jarosch et al., 2021). However, much of the existing research has been limited by data availability, focusing on specific occupations (Mas and Moretti, 2009) or experimental settings (Falk and Ichino, 2006). Few studies have simultaneously analyzed worker and peer quality in the same framework. For example, following the approach of Arcidiacono et al. (2012), recent papers have linked coworker effects to wages (Cornelissen et al., 2017; Battisti, 2017; Cardoso et al., 2022; Messina et al., 2023). However, these findings often vary by country, team definitions, and identification strategies[1]. This paper advances the literature by being the first to focus explicitly on the peer quality gap, providing a novel perspective on how relative differences between workers and their peers shape workplace outcomes.

In addition, our research contributes to the understanding of worker mobility determinants. Workers often accept lower wage jobs to then move and improve their situation, seeking higher wages, better wage growth trajectories (Postel-Vinay and Robin, 2002), improved occupational matches (Groes et al., 2015), , or opportunities to acquire skills and knowledge for career progression (Gathmann and Schönberg, 2010; Cortes and Gallipoli, 2018). Our paper bridges the literatures on peer effects and worker mobility by demonstrating how the peer quality gap influences both wages and job transitions, including movements between employment and unemployment. By linking these two dimensions, we provide novel evidence on how relative peer quality shapes labor market dynamics. Peer pressure emerges as a likely mechanism driving these behaviors, highlighting the interplay between social dynamics and economic incentives.

Our second contribution adds to the worker-to-firm allocation literature, particularly by advancing our understanding of how sorting mechanisms influence wage dispersion. A key innovation of the AKM model (Abowd et al., 1999), is its ability to decompose

[1] Table A2 summarizes the results from the literature for various fix effects specifications and peer groups definitions.

wage dispersion, with match quality dispersion identified as a critical component[2].The observation that high-quality individuals often sort into high-quality firms is puzzling[3], as improving allocation could have substantial macroeconomic benefits. Several frameworks have been proposed to explain sorting[4], including market power, firm size, risk aversion, and multidimensional mismatch. For instance, recent research by Eliason et al. (2023) shows that while social connections influence hiring, they do not necessarily increase assortativity. Although theoretical insights into allocation mechanisms have advanced (Galichon, 2018; Lindenlaub, 2017; Lindenlaub and Postel-Vinay, 2023; Tan, 2023), empirical evidence on the underlying drivers remains scarce. Our results contribute to this literature by demonstrating how the peer quality gap evolves over workers' careers. We show that firms may exploit heterogeneous effects on high- and low-quality workers in high-skill jobs to decrease overall worker quality and maximize profits, shedding light on a novel mechanism in worker allocation.

Our third contribution is methodological. This paper addresses a key gap in the literature by presenting a feasible method to estimate heterogeneous peer effects across both observable and unobservable characteristics. Calculating heterogeneous effects in large social networks is uncommon, as peer effects are often assumed to be homogeneous (Comola et al., 2022). However, if workers group based on shared characteristics, they are also likely to exhibit similar behaviors, making it critical to account for heterogeneity. To the best of our knowledge, only a few papers have explored peer effect heterogeneity in the workplace. Messina et al. (2023) documented within and between-gender heterogeneous effects for one large city in Brazil using an empirical approach. In contrast, Jarosch et al. (2021) developed and estimated a structural model in which learning from coworkers has long-term effects on wages and career trajectories. Our approach builds on these studies by offering a practical statistical framework to analyze peer effects across diverse populations, uncovering dynamics that have been previously overlooked in the literature.

Outline The remainder of the paper is organized as follows. Section (2) describes our main data sources and presents descriptive statistics in the estimation sample used. Section (3) details the empirical design and the methods to identify heterogeneous effects for observable and unobservable characteristics. Section (4) present the results, and Section (5) concludes.

2 Data

This section outlines the primary data source and the variables used. We will also define our sample, peer groups, and worker career trajectories.

Quadros de Pessoal. We use Quadros de Pessoal (QP), a Portuguese linked-employer-employee-dataset. Nowadays, QP captures information on over 350,000 firms and 3 million workers. The dataset enables unique identification of each worker and firm, facilitating the tracking of worker careers. It provides detailed information on remuneration, job

[2]The quantitative importance has been in the center of discussion lately (Bonhomme et al., 2019; Andrews et al., 2008; Kline et al., 2020)

[3]Several studies have shown how firms can influence wage dispersion through changes in firm quality dispersion (Card et al., 2013; Song et al., 2019; Leitao et al., 2022), and how differentiated sorting can contribute to wage inequality (Card et al., 2016).

[4]Eeckhout (2018) provides a comprehensive review of the different approaches in the sorting literature.

conditions, and job transitions. While most linked employer-employee datasets originate from social security administrative records, QP is directly reported by employers as a tool for monitoring and controlling worker labor conditions. All private firms in Portugal with at least one employee are required to submit their active personnel information to the Ministry of Employment each October. QP includes detailed information on the worker, the firm, and the *match*. Firms report demographic characteristics such as age, gender, and education, along with firm characteristics like establishment, economic activity, size, and sales. Each registry also characterizes the work relationship, providing information on remuneration, hours worked, firm tenure, occupation, and the collective agreement that limits the work relationship. Following recent literature (Cornelissen et al., 2017; Cardoso et al., 2022), we define peer groups within the dataset based on the principle that peer workers share the same job and working conditions. We define a peer group as all non-singleton workers[5] sharing the same occupation and collective agreement within the same establishment in a given year.

Sample. Our sample spans from 1994 to 2021. We harmonize sector, occupation, education, and establishment identifiers across the panel. We retain complete and valid observations from full-time workers, excluding observations with missing information in key variables of interest or abnormal values in wage, hours, and tenure. The sample comprises unique worker-year observations with the dominant employer. We restrict our focus to workers aged 16 to 61 earning more than half of the national minimum wage for that year. We exclude workers from sectors with cyclical employment patterns and high rates of informality, such as agriculture, fishing, and construction. We impose three restrictions for accurate identification of peer fixed effects. Following Cardoso et al. (2022), or computational feasibility, we initially limit establishment size to a maximum of 2000 workers. Second, we restrict the sample to non-singleton peer groups (occupation $\times$ collective agreement $\times$ establishment $\times$ year). Lastly, we ensure that the remaining peer groups are connected by worker mobility – *largest connected set* –(Abowd et al., 2002; Kline et al., 2020). Table A1 summarizes the impact of these restrictions on the data to arrive at our final sample.[6]

The remaining sample follows the careers of $3,969,261$ workers across $274,396$ firms. Career histories are highly detailed, encompassing granular information on the establishment, occupation, collective agreements, and wages for each of the $24,046,313$ worker–year observations. Furthermore, the unique identifiers enable us to identify career transitions, including interruptions, labor market exits, and changes between firms and peer groups. This allows us to observe both internal and external job-to-job transitions. Log-hourly real wage is our preferred measure of labor income. A worker's total earnings consists of base wage, overtime payments, and subsidies (e.g., meals, tenure rewards, and ordinary and extraordinary bonuses). Total earnings are deflated using the monthly CPI for October of each year. To account for variations in hours, we divide real total earnings by the total reported hours. Total hours are the sum of normal and extra hours. We apply the logarithm to this final measure. Two factors ensure the quality of the earnings data: first, reporting by the employer guarantees it is not underestimated. Second, institutional mechanisms, such as continuous inspections by the Ministry of Labor and the availability of information to any worker in the firm, make it a valuable tool for internal HR and used by unions for negotiating collective agreements.

The data also includes the identifier for the collective agreement governing the work contract. Given that each collective agreement sets minimum wages and benefits by job

[5]We observe singleton workers only once.

[6]See Appendix **??** for a detailed description of the data cleaning.

Table 1: Descriptive Statistics

	Statistic				
	Mean	St. Dev.	Min	Max	Nº Unique
Worker-level					
Log of houly wage	1.82	0.55	-0.06	8.94	
Male	0.51	0.50	0.00	1.00	
Age (years)	37.77	10.74	16.00	64.00	
Tenure (months)	100.50	107.78	0.00	600.00	
Log of firm size	5.09	2.07	0.69	10.13	
Log of estab. size	4.27	1.55	0.69	7.60	
Establishment-level					
Number of groups	1.53	1.38	1.00	41.00	
Log of estab. size	2.42	1.11	0.69	7.60	
Firm-level					
Number of estab.	1.89	10.53	1.00	1038.00	
Log of firm size	2.58	1.24	0.69	10.13	
Worker-level Fluxes					
Promotion	0.08	0.28	0.00	1.00	
Separation	0.23	0.42	0.00	1.00	
Switch jobs	0.08	0.27	0.00	1.00	
Hiring	0.28	0.45	0.00	1.00	
Hiring from unemp.	0.09	0.29	0.00	1.00	
Distinct values of:					
Observations					24046313
Workers					3969261
Peer groups					2227467
Establishments					404359
Firms					274396
Job titles					33717
Occupations					110
Years					27

Source: Quadros de Pessoal, MTSSS, 1994 − 21.

Note: This table shows the main descriptive statistics of our sample. The table has 5 sections. The first section reports statistics at the worker-level. The establishment- and firm-level sections report statistics from collapsed data containing only distinct combinations between years and establishments/firms, respectively. The forth section contains variables about worker fluxes: promoted refers to workers that have increased their professional category inside a firm since the previous year; a worker separates whenever he's employed at a firm where he won't be employed in the following year; switching jobs consists of a separation to another firm and to be a robust job switch a worker must have been employed for at least 3 years at the firm from which he separated from; lastly, a hiring is a worker that was not at that firm in the previous year and a hiring from unemployed is a worker a hire of a worker that has been *out of sample* for at least three years. The last section reports the distinct values of certain categories.

category within the firm, interacting the collective agreement with the occupation proxies the set of tasks and the skill requirements relative to a wage floor. In our case, the interaction of collective agreement with occupation provides valuable information for identifying the worker's peer group, allowing us to group worker peers with the same job content and work conditions. Interacting it further with time and establishment ensures physical proximity. This ability to follow worker and team composition enables us to separately study how workers utilize changes in peer groups to advance their careers and how firms select the composition of the peer groups. The final panel has 27 years, includes $404,359$ establishments, and comprises $33,717$ job-titles. This results in $2,227,467$ peer groups.

Table 1 presents the main descriptive statistics of our sample, covering variables at the worker, establishment, and firm levels. The table reflects Portugal's predominance of small firms, where the average firm employs 11 workers. Moreover, the majority of firms operates a single team (71%) and only 7% have more than one establishment. However, since larger firms employ a greater share of the workforce, most workers are employed at a multi-team firm (77%).

Even though Table 1 reports an almost equal gender composition in our sample, this is the result of a male majority in the 1990s, followed by a female majority around 2010, and an equal representation in recent years. Moreover, this gender shift was not uniform across age groups. As in many European countries, Portugal's workforce is aging (Aiyar and Ebeke, 2016). Figure B2 shows that while both men and women are aging, the trend is particularly pronounced among women. This increase is not due to a lower entry of young females but rather because the share of older women was very small in the 1990s, reflecting the low female participation rates in previous decades. It's plausible that these different demographic groups may respond differently to average peer ability, potentially influencing our estimates of peer effects.

3 Methods

The primary aim of this study is to quantify how much of workers' wages stems from the gap between their own quality and the average quality of their peers. Three primary causal channels may account for these gap effects: knowledge spillovers, contagious enthusiasm/malaise and peer pressure (Mas and Moretti, 2009). Knowledge spillovers occur when one worker's know-how benefits another, increasing marginal productivity, and hence also effort and output. Although undeniably important in some specific industries (Ichinowski et al., 1997), we do not expect this to be the driver in most scenarios (Cornelissen et al., 2017; Mas and Moretti, 2009), also because they imply that better workers can be made less productive via negative knowledge spillovers from worse workers. Contagious enthusiasm, by contrast, increases the motivation of workers who are worse, also increasing their effort (Brune et al., 2022). Lastly, peer pressure occurs when a worker feels discomfort working less diligently than others. This pressure can result from formal actions, such as reports to management, or informal reactions, like social ostracism. Thus, through peer pressure, a better worker can incentivize effort and a worse worker can disincentivize it (Kandel and Lazear, 1992).

To capture gap effects we cannot simply use the difference between the average peer wage and own wage. This econometric challenge is known as the reflection problem (Manski, 1993; Angrist, 2014). Because a worker's wage is shaped by their peers, whose wages are simultaneously influenced by that worker, we get an intertwined feedback loop.

Diagram B1 provides a simple example of a team consisting of two workers, A and B. The wage of worker B is influenced by the quality of worker A. Consequently, we would face difficulties in distinguishing whether B's wage is causing A's wage or merely reflecting it. To avoid the reflection problem, we estimate peer effects based on quality rather than wages. However, because quality is an unobserved variable, our model cannot be estimated using OLS. Therefore, we adapt an algorithm developed by Arcidiacono et al. (2012) to estimate both quality and the quality gap effect simultaneously.

In line with the literature, we measure average peer quality by averaging the worker fixed effects of a given worker's colleagues (Cornelissen et al., 2017; Battisti, 2017; Cardoso et al., 2022; Hong and Lattanzio, 2022). However, we modify how peer quality enters our equation by using the *quality gap* instead. The quality gap refers to the difference between a worker's quality and the respective average peer quality. Our approach offers three main advantages, even though, econometrically, only the worker FEs are re-scaled, with no change in model fit. Firstly, it provides a more accurate estimation of quality by aligning our econometric structure with the main mechanisms behind peer effects. For example, the peer pressure literature models it as a gap (Mas and Moretti, 2009; Kandel and Lazear, 1992). Even in Cornelissen et al. (2017), the quality gap is the theoretical mechanism through which peer quality influences colleagues, even thought they later estimate peer effects without using the quality gap. Second, because the interpretation of gap effects as an elasticity between peer quality and wages is merely an approximation, the re-centering of average peer quality, by reducing the magnitudes of the peer effects, it will impact this interpretation used in some previous studies (Cornelissen et al., 2017; Hong and Lattanzio, 2022; Cardoso et al., 2022). However, the (correct) interpretation originally used by Arcidiacono et al. (2012) yields similar interpretations. Lastly, it enables us to directly study worker mobility and team composition by closely aligning with the mechanisms behind these dynamics.

Besides estimating peer effects, we extend the model in two ways. The first extension introduces response heterogeneity, allowing different groups of workers to have different levels of responsiveness to peer quality. In other words, we can group workers by observable characteristics and retrieve a peer gap effect for each of the groups, which allows us to estimate how the quality gap affects different demographic groups. If a group of workers is less affected by peer quality, it will benefit less from better peers but also lose less from worse ones. This tool allows us to identify which groups are driving the overall quality gap effects. This first extension only allows the heterogeneity to be analyzed with observed variables. The analysis is restricted to observable variables, which may be at the worker level (e.g., age, gender), distinguishing workers within a team, or at the team level or higher (e.g., location, occupation), comparing workers between teams. This approach is the only method that allows for such an analysis of response heterogeneity based on observable characteristics (Messina et al., 2023; Comola et al., 2022). To estimate the simple peer effects, we use the algorithm of Arcidiacono et al. (2012) with the efficiency improvements made by Cardoso et al. (2022). Estimating heterogeneous peer effects requires only slight modifications to this algorithm, as also developed by Arcidiacono et al. (2012).

The second extension incorporates endogenous groupings to capture response heterogeneity, basing groupings on model-estimated variables rather than solely on observable variables. Our primary focus is on quality, intending to split each team into two categories: workers outperforming their peers and those performing below the team average. This approach can also be combined with observable heterogeneity, such as skill level, providing us a different effect of quality gap for the better and worse workers by skill level.

A central assumption of basic peer effects models is that team composition functions as a zero-sum game: total wages remain constant as one worker's gain translates to another's loss. Consequently, any arrangement of workers is equally effective concerning overall wages, making the effects of team composition purely distributional. This extension allows us to challenge this restrictive assumption and assess whether the benefits and costs of peer interactions balance or if one prevails. Since we anticipate that quality gap effects will favor workers below average while disadvantaging those above, we examine whether worse workers gain more than better ones lose (or vice versa). This allows us to answer to a critical question regarding team composition: Are high-variance teams more beneficial than homogeneous ones, or is a more balanced team preferable? Under what circumstances might one composition be more advantageous than the other? If gap effects are bigger for workers worse than their peers, it implies that they benefit more from the higher-quality peers than those better workers have to lose. Thus, a firm seeking to maximize overall effort should avoid homogeneous teams. On the other hand, if better workers are more negatively impacted by the quality gap, it indicates that a firm should aim for low-variance teams, as the better workers' losses outweigh the benefits they provide to lower-quality workers. Being us able to estimate different effects across workers of different quality, allows us to understand how firms optimize team composition.

3.1 Effects from peer quality gap

Consider a peer group p that employs n_p workers i, comprising the set of workers $\mathcal{I}_p$. There are M observations of N distinct workers working in P peer groups. To capture the effect of the quality gap on wages, we estimate the following baseline wage equation:

$$log(y_{it}) = \alpha_i + \eta \left[\bar{\alpha}_{p(i,t)\sim i} - \alpha_i\right] + \boldsymbol{X}'_{it}\boldsymbol{\beta} + \varepsilon_{it}, \tag{1}$$

where y_{it} is the real hourly wage, X_{it} is a vector of controls (including fixed effects), α_i is the fixed effect of worker i, $\bar{\alpha}_{p(i,t)\sim i}$ is the average worker fixed effects of the peers of worker i, and ε_{it} is the residual. Model 1 can be viewed as an AKM model (Abowd et al., 1999) with peer effects, since we estimate it with age, tenure, interaction of collective agreement with occupation fixed effects, and firm fixed effects in X_{it}. To clarify further, the average peer quality is calculated while excluding the worker's own quality:

$$\bar{\alpha}_{p(i,t)\sim i} = (n_p - 1)^{-1} \sum_{j \in \mathcal{I}_p \sim i} \alpha_j.$$

We'll refer to $\alpha_i - \bar{\alpha}_{p(i,t)\sim i}$ as the quality gap, which enters 1 negatively. Thus, a one standard deviation decrease in the quality gap, which implies a peer quality improvement, increases wages by $(e^{\eta \times sd} - 1) \times 100\%$.[7] Assuming $\hat{\eta} > 0$, three cases occur: (i) A below-average worker gains from higher peer quality; (ii) A worker of average quality sees no effect; (iii) An above-average worker experiences a wage reduction due to lower peer quality.

We now rewrite model 1 using matrix notation:

$$\boldsymbol{Y} = \boldsymbol{D\alpha} + \eta \left[\boldsymbol{W} - \boldsymbol{I}\right] \boldsymbol{D\alpha} + \boldsymbol{X\beta} + \boldsymbol{e}, \tag{2}$$

where $\boldsymbol{Y}_{(M\times 1)}$ is a vector with the log real hourly wages, $\boldsymbol{X}_{(M\times k)}$ contains k control variables and fixed effects, $\beta_{(k\times 1)}$ is a vector of parameters, $\alpha_{(N\times 1)}$ is a vector containing the worker

[7] ADD A COMPREHENSIVE FOOTNOTE of INTERPRETATION.

fixed effects, $\boldsymbol{D}_{(M\times N)}$ is the worker design matrix, and $\boldsymbol{W}_{(M\times M)}$ is a symmetric block diagonal weighting matrix, defined as:

$$\boldsymbol{W} = diag(\mathcal{W}), \text{ where the set } \mathcal{W} = \{\boldsymbol{w}_p \,\forall\, p \in [1;P]\}.$$

The sub-matrices $\boldsymbol{w}_p$ concern one peer group each, and are defined as the following:

$$\boldsymbol{w}_p = \frac{\mathbb{1}_{n_p}\mathbb{1}'_{n_p} - \boldsymbol{I}_{n_p}}{n_p - 1},$$

where $\mathbb{1}_{n_p}$ is a vector filled with 1s of size n_p.

We follow closely Cardoso et al. (2022) in our estimation method. To find the parameters of interest, as usual, we proceed to minimize the sum of square residuals:

$$\arg\min_{\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\eta}} \boldsymbol{e}'\boldsymbol{e} = \{\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\eta}}\},$$

which leads to the following first order conditions:

$$\frac{\partial L(\cdot)}{\partial \hat{\boldsymbol{\alpha}}} = \nabla_{\boldsymbol{\alpha}} L(\cdot) = \vec{0} \;\wedge\; \frac{\partial L(\cdot)}{\partial \hat{\boldsymbol{\beta}}} = \nabla_{\boldsymbol{\beta}} L(\cdot) = \vec{0} \;\wedge\; \frac{\partial L(\cdot)}{\partial \hat{\boldsymbol{\eta}}} = \nabla_{\boldsymbol{\eta}} L(\cdot) = \vec{0}.$$

The conditions for $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\eta}}$ can be solved by OLS-like algorithms designed to deal high-dimensionality (Correia et al., 2016). However, the condition for $\hat{\boldsymbol{\alpha}}$ simplifies to:

$$\mathbf{D}'\tilde{\mathbf{W}}\tilde{\mathbf{W}}\mathbf{D}\hat{\boldsymbol{\alpha}} = \mathbf{D}'\tilde{\mathbf{W}}[\mathbf{Y} - \mathbf{X}\tilde{\boldsymbol{\beta}}], \tag{3}$$

where $\tilde{\mathbf{W}} = \hat{\eta}[\mathbf{W} - \mathbf{I}] + \mathbf{I}$. Because of the size of the $\mathbf{D}'\tilde{\mathbf{W}}\tilde{\mathbf{W}}\mathbf{D}$ matrix, we cannot simply invert it and solve this equation. Therefore, we solve for $\hat{\boldsymbol{\alpha}}$ using the conjugate gradient method. However, because we are solving the first order conditions each at a time, we must iterate between them until convergence is reached, i.e., no coefficient substantially changes with one more iteration.

3.2 Peer quality Gap Effects Across Groups

In this section we extended our econometric specification to estimate heterogeneous quality gap effects across different groups, i.e., allowing some groups of workers to be more or less affected by peer quality. A worker that is part of a group that is more responsive to peer quality gaps will benefit more from better peers but also lose more from worse ones, compared to workers in less responsive groups.

Consider the diagonal matrix $\boldsymbol{Q}_q$, of size $M \times M$, where the entries along the diagonal represent a binary vector (or dummy variable):

$$\boldsymbol{Q}_q = \text{diag}(\mathbb{1}_{i\in q}).$$

This vector indicates whether each observation belongs to group q. Specifically, for an observation i, the i-th diagonal entry of $\boldsymbol{Q}_q$ is 1 if the observation belongs to group q and 0 otherwise:

$$\mathbf{1}_{i\in q} = \begin{cases} 1, & \text{if observation } i \text{ belongs to group } q \\ 0, & \text{otherwise} \end{cases}.$$

All off-diagonal entries in $\boldsymbol{Q}_q$ are zero, making $\boldsymbol{Q}_q$ a sparse matrix. For example, if there are 3 groups q, we get Q_0, Q_1 and Q_2, filled with 1's and 0's in their diagonal. The allocation into groups must be mutually exclusive and identifiable. Mutually exclusive because each observation can only belong to one group q. Identifiable because each observation must belong to a group based on an observable variable. Thus, $\sum_q \boldsymbol{Q}_q = \boldsymbol{I}$ holds by definition. Thus, the heterogeneous extended model can be written as:

$$\boldsymbol{Y} = \boldsymbol{D}\boldsymbol{\alpha} + \sum_q \eta_q \boldsymbol{Q}_q \left[\boldsymbol{W} - \boldsymbol{I}\right] \boldsymbol{D}\boldsymbol{\alpha} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}, \tag{4}$$

where each η_q is the gap quality effect for each heterogeneity group q. The interpretation of $\hat{\eta}_q$ is the same as in the simple case, but now it's referring only to observations of group q.

Compared to the simple case, all first order conditions remain the same, except for the condition of $\boldsymbol{\alpha}$. Still, the only necessary change in the condition of $\boldsymbol{\alpha}$, shown in equation 3. is in the $\tilde{\mathbf{W}}$ matrix, that is now defined as:

$$\tilde{\mathbf{W}} = \sum_q \hat{\eta}_q \boldsymbol{Q}_q [\mathbf{W} - \mathbf{I}] + \mathbf{I}.$$

Furthermore, we'll also employ a different method to solve the condition of $\boldsymbol{\alpha}$. Instead of the conjugate gradient, we'll instead use the biconjugate gradient method (CITATION NEEDED). [8]

3.3 Different Gap Effects for Better and Worse Workers

When discussing heterogeneous effects across different groups, a natural question is whether there are different effects based on worker quality. What are the impacts of the quality gap for "better" and "worse" workers? Grouping based on worker quality α_i presents a challenge since it is unobserved.

In this second extension, we allow the quality gap effect to vary based on a variable estimated by the model, rather than solely relying on observable characteristics. Conceptually, this approach is similar to model 4. However, the difference lies in the classification of workers as belonging group q, which is now a function of $\hat{\boldsymbol{\alpha}}$. This means that the definition of group q depends on the distribution of the estimated parameter. While $\boldsymbol{Q}_q$ in model 4 was given, we now allow it to be a function of the worker fixed effects: $\hat{\boldsymbol{Q}}_q(\hat{\boldsymbol{\alpha}})$.

This approach enables us to estimate different effects of the quality gap for workers who perform either better or worse than their team average, as the quality gap is solely determined by $\hat{\boldsymbol{\alpha}}$. Importantly, this model allows for endogenous response variquality as well as standard heterogeneity (discussed in Section 3.2). To address both types of heterogeneity, we define the parameters as follows: η_{qw} for workers in group q who perform below their team average ($\bar{\alpha}_{p(i,t)\sim i} - \alpha_i \geq 0$), and η_{qb} for workers in group q who perform above the team average.

Although we can classify workers as simply "better" or "worse", our model will not have a simple dummy for each category, because that would not yield a unique and stable solution. This is due to the non-convex nature of the optimization problem; even a small increase in α_i may shift worker i from below to above the team average, resulting in a switch of the coefficient from η_{qw} to η_{qb}. To address this non-convex problem, we approximate it with a

[8] While the conjugate gradient method requires the matrix $\mathbf{D}'\tilde{\mathbf{W}}\tilde{\mathbf{W}}\mathbf{D}$ to be symmetric and positive-definite, which is the case in the simple estimation, with heterogeneous effects it's not necessarily symmetric anymore. Thus, we use the biconjugate method, which is a more general method that does not require such assumptions.

smoothing function, specifically the standard logistic function $\sigma(\cdot)$, which is widely used for such approximations (CITATION NEEDED).

Therefore, instead of directly inserting in our model a dummy variable indicating "better" or "worse", achieved by noting if $\bar{\alpha}_{p(i,t)\sim i} - \alpha_i$ is either positive or negative, we instead insert $\sigma(\bar{\alpha}_{p(i,t)\sim i} - \alpha_i)$, where $\sigma(\cdot)$ is the standard logistic, defined has:

$$\sigma(\bar{\alpha}_{p(i,t)\sim i} - \alpha_i) = \frac{1}{1 + e^{-\bar{\alpha}_{p(i,t)\sim i} + \alpha_i}}.$$

We apply the logistic function to every observation:

$$\boldsymbol{\sigma} = \sigma([\boldsymbol{W} - \boldsymbol{I}]\,\boldsymbol{D}\boldsymbol{\alpha}) = \begin{bmatrix} \sigma(\bar{\alpha}_{p(1,t)\sim 1} - \alpha_1) \\ \sigma(\bar{\alpha}_{p(2,t)\sim 2} - \alpha_2) \\ \cdots \\ \sigma(\bar{\alpha}_{p(N,t)\sim N} - \alpha_N) \end{bmatrix},$$

in order to get $\boldsymbol{\sigma}$, a continuous vector bounded between 0 and 1. A "better" worker is above 0.5 and a worse worker is below, but this smoothed indication also reflects intensity. We can now express the complete model as follows:

$$\boldsymbol{Y} = \boldsymbol{D}\boldsymbol{\alpha} + \sum_q \eta_{dq} diag(\boldsymbol{\sigma})\boldsymbol{G}_q\boldsymbol{\alpha} + \sum_q \eta_{bq}\boldsymbol{G}_q\boldsymbol{\alpha} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}, \tag{5}$$

where $\boldsymbol{Q}_q\,[\boldsymbol{W} - \boldsymbol{I}]\,\boldsymbol{D} = \boldsymbol{G}_q$ and $\eta_{dq} = \eta_{qw} - \eta_{qb}$.

Because the vector $\boldsymbol{\sigma}$ depends on $\boldsymbol{\alpha}$, the first order condition of $\hat{\boldsymbol{\alpha}}$ from model 5 changes. The new gradient with respect to $\hat{\boldsymbol{\alpha}}$ is given by:

$$\nabla_\alpha \boldsymbol{e}'\boldsymbol{e} = -2\boldsymbol{D}'\hat{\boldsymbol{e}} - 2\sum_q \left\{ \boldsymbol{G}_q' \Big[\hat{\eta}_{dq} diag(\hat{\boldsymbol{\sigma}} + (\boldsymbol{G}_q\hat{\boldsymbol{\alpha}}) \odot \hat{\boldsymbol{\varsigma}}) + \hat{\eta}_{bq} \Big] \right\} \hat{\boldsymbol{e}}$$

where $\boldsymbol{\varsigma}$ is the derivative of the logistic function with respect to the quality gap, defined as:

$$\frac{\partial \boldsymbol{\sigma}}{\partial\,[\boldsymbol{W} - \boldsymbol{I}]\,\boldsymbol{D}\boldsymbol{\alpha}} = diag(\boldsymbol{\sigma} \odot [\mathbb{1} - \boldsymbol{\sigma}]) = diag(\boldsymbol{\varsigma})$$

The resulting matrix is a $(M \times M)$ Jacobian diagonal matrix (because all $\hat{\alpha}_i$s are independent between them); $\odot$ is the Hadamard product (element wise); and $\mathbb{1}$ is a vector of 1s of equal size. Since the gradient is non-linear, we'll use the gradient descent method to solve it, with each iteration h, updating the parameters as follows:

$$\tilde{\boldsymbol{\alpha}}_{h+1} = \tilde{\boldsymbol{\alpha}}_h - \gamma\nabla_{\alpha,h}\boldsymbol{e}'\boldsymbol{e}$$

where γ is the learning rate. We alternate between all first order conditions until convergence is achieved, i.e., until the $\hat{\boldsymbol{\eta}}$ parameters change very little between iterations.

4 Results

This section presents our findings on how quality gaps influence labor earnings, their persistence over time, and their implications for worker mobility and team composition. Given the direct relationship between quality gaps and peer effects, we start describing the identification source and the advantages of using the recentered version. We first benchmark our results, comparing them with the existing literature. We explore the effects of quality gaps on labor earnings and their persistence. Subsequently, this section leverages the frameworks presented in sections 3.2 and 3.3 to examine heterogeneous effects across various groups in our sample. The methodology enables differentiation of heterogeneous effects based on observable characteristics and the level of worker quality, a constant unobserved characteristic. Compared to peer effects, the quality gap's primary advantage lies in its immediate connection to the dynamics of employment and job mobility.

How does the quality gap affect workers' and firms' behavior? We start by analyzing the implications for workers. When the quality gap is substantial, indicating that a worker's quality significantly exceeds the average quality of their peers, model (1) implies a wage penalty if $\hat{\eta} > 0$. If the worker has an outside option on a similar job but with a quality peer gap reduction, that option would likely be preferred. A substantial quality gap, therefore, could induce mobility[9]. From the firm's perspective, a large quality gap might increase profits, through an increase in total effort, when the worker quality dispersion within the team effects on better and worse workers are heterogeneous. Altogether, as firms incorporate worker behavior, they optimize between peer team composition and the worker exit rate. We analyze this novel aspect, linking quality gaps and worker mobility, from both sides of the market. If the effect is sufficiently strong, it could influence the composition of peer groups and influence the allocation of workers to firms.

4.1 The effect of quality gap on present and future labor earnings

Leveraging on employer–employee data and following our peer group definition (job-title $\times$ year $\times$ establishment), we examine the impact of the quality gap on current and future wages. This section explores how the peer quality gap compares to previous literature, which reports on the magnitude and significance of peer fixed effects on wages (Cardoso et al., 2022; Cornelissen et al., 2017; Battisti, 2017), and their time persistence (Hong and Lattanzio, 2022).

Table 2 presents the results of estimating equation 2 in our sample. Our preferred model controls for unobserved constant heterogeneity at the worker, firm, the interaction of collective agreement with occupation, and year levels. The value of $\hat{\eta}$, the parameter that captures the quality gap, is 0.3547. This value indicates an increase in labor earnings of approximately 12.8% if the worker's quality gap increases 1 s.d.[10] Thus, the place of a worker in the distribution of peer quality gaps has a substantial and significant effect on wages. Moreover, the effect on wages from a 1 s.d. increase in the gap is equivalent to 22% of the effect of a 1 s.d. increase in individual quality.[11] A third possible way of interpreting $\hat{\eta}$ is as an approximation to an elasticity.[12] Thus, 10% increase in peer quality leads to an increase

[9]A substantial quality gap could also generate an increase in compensation within the current firm, e.g. promotion. We do not study this phenomenon due to data constraints.

[10]The standard deviation of the peer quality gap fixed effect is 0.3387, and the standard deviation of worker quality is 0.4585.

[11] $\frac{e^{0.3387\times0.355}-1}{e^{0.4585}-1} = 22\%$.

[12]While η is not an elasticity due to the model functional form, it does have an elasticity approximation

in wages of approximately 3.5%. To put it in perspective, a worker passing from a null quality gap to a 1 s.d. positive quality gap is equivalent to passing from a firm at the 38 percentile of the firm fixed effects distribution to a firm on the 67 percentile. WRITE SENTENCE 1 The large and significant effect of the peer quality gap on labor income supports the notion that it may trigger behavioral responses in workers, motivating our focus on worker mobility.

Table 2: Preferred Model of Peer Gap Effects

Dependent Variable:	Log of hourly real wage
Model:	(1)
Variables	
$\hat{\bar{\alpha}}_{-i} - \hat{\alpha}_i$	0.3547***
	(0.0007)
age	0.0307***
	(0.0000)
age^2	-0.0002***
	(0.0000)
tenure	0.0009***
	(0.0000)
tenure^2	0.0000***
	(0.0000)
Fixed-effects	
worker	✓
firm	✓
job title	✓
year	✓
Fit statistics	
Observations	24,046,313
R^2	0.90598
Within R^2	0.62525

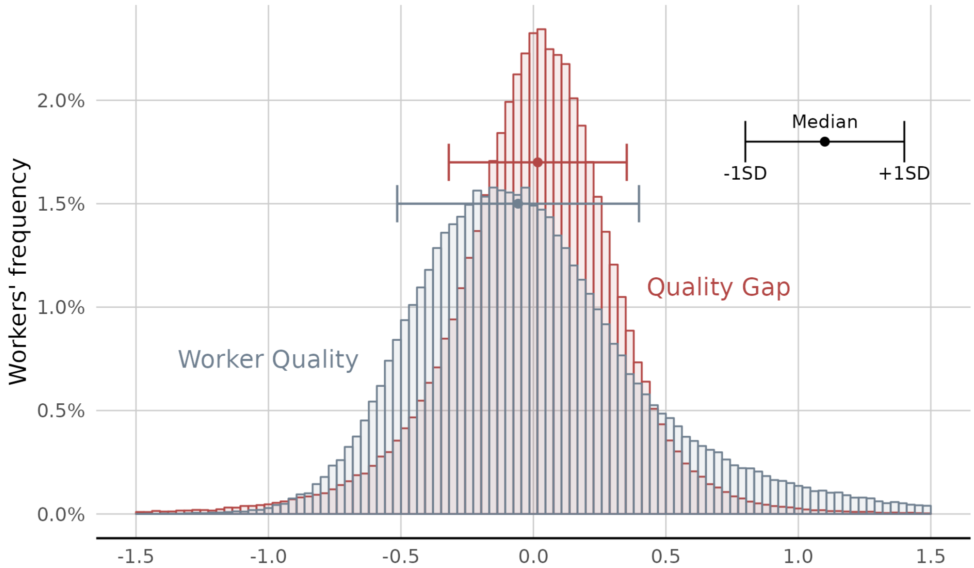
Source: Quadros de Pessoal, MTSSS, $1994 - 21$.
Note: This table shows the coefficients and fit statistics from our preferred model 1. Clustered (peer group) standard-errors in parentheses. Significance Codes: ***: 0.01, **: 0.05, *: 0.1

To compare our proposed measure of the peer quality gap with the peer fixed effects, we present Table A3. While the coefficients may initially appear different, the effects on wages are similar across the regressions.[13] This is not surprising as the variation exploited for identification is the same in both cases. It relies on a combination of variation in worker quality, peer group size, and peer quality. In our preferred specification, we focus on the variation in group size and exploit the variation induced by workers leaving or entering the peer group (as in Cornelissen et al. (2017)). Our results align with recent research using the same Portuguese data (Cardoso et al., 2022). The difference in coefficients arises from a rescaling of the person effects, resulting in the marginal effect of wages on the person effect being a function of the peer quality gap. Column (4), in Table A3, presents our preferred specification, which includes firm, job-title, and year fixed effects. The level of disaggregation proposed in the fixed effects aligns with the empirical design of the research question. In this

by assuming peer homogeneity. See section **??** for a discussion on the interpretation of the peer gap quality parameter.

[13] ADD COMPARISON FOR THE PREFERRED SPECIFICATION

Figure 1: Distribution of Worker Quality and the Quality Gap



Source: Quadros de Pessoal, MTSSS, $1994-21$.
Note: The figure illustrates, in red, the distribution of worker quality, quantified through worker fixed effects as derived from model 1. In contrast, the blue curve represents the quality gap, defined as the disparity between a worker's own fixed effect and the average worker fixed effect of their peer group. For visualization purposes, only a 1% sample of the data is depicted. Both distributions are accompanied by a bandwidth that highlights the median (the central value), as well as one standard deviation above and below this median.

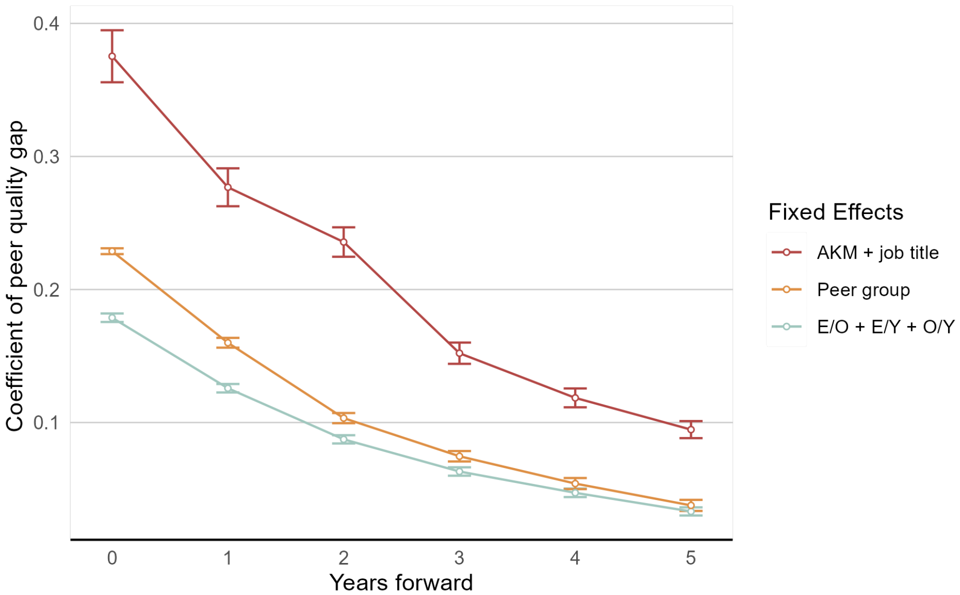
article, we do not employ a fully year-interaction model (Hong and Lattanzio, 2022; Messina et al., 2023), or the year-peer group fixed effect (Cardoso et al., 2022). This choice is justified by our ultimate goal of understanding worker mobility, which is entirely absorbed when interacting with year-fixed effects. Additionally, our specification normalizes for possible positive assortative matching or cases where worker quality is highly correlated with specific occupations or establishments. This approach avoids the potential for incidental parameter bias that might arise from interacting with the peer group and the year.

To investigate the impact of the quality gap on future wages we follow Hong and Lattanzio (2022)'s empirical strategy.[14] Figure 2 illustrates the effect of the peer quality gap on future wages for different fixed effects specifications. Our findings reveal that the peer quality gap is persistent over time but eventually dissipates. For our preferred specification (red line), the influence of the quality gap on wages declines from $12.8\,\%$ to $10.13\,\%$ after one year. This effect gradually decreases further to $8.28\,\%$, $5.4\,\%$, and ultimately to around $3.3\,\%$ after five years. Given that the peer quality gap influences both current and future wages, it represents a substantial cost for workers. This high cost might influence mobility decisions, as workers may consider outside options that could lead to job-to-job mobility (Postel–Vinay and Robin, 2002).

This section discussed how our modeling choices led to these estimates and how they relate to the existing literature. Our main contribution lies in our ability to isolate these effects for specific groups and establish their connection to worker mobility choices, which

[14] As them, we use future wages, to assess the persistence. We test different versions of the model to capture mobility, and also more restrictive versions in which identification is attained by group size and composition change. We estimate the model $log(y_{i,t+h}) = \alpha_i + \eta \left[\bar{\alpha}_{p(i,t)\sim i} - \alpha_i\right] + \boldsymbol{X}'_{it}\boldsymbol{\beta} + \varepsilon_{it}$, where h is the number of periods ahead. Furthermore, we restrict each lead sample $t+h$ to their respective large connected set.

Figure 2: Quality Gap Effects on Future Wages



Source: Quadros de Pessoal, MTSSS, $1986-21$.

Note: The figure illustrates the effect of the peer quality gap on future wages. The coefficients presented in the figure correspond to various η coefficients derived from model 1, applied across different samples and adjusted with varying fixed effects. Rather than estimating the model with $log(y_{it})$, it is re-estimated using $log(y_{i,t+h})$, where h represents an integer ranging from 0 to 5. Changing the dependent variable alters the sample size, as not all workers remain in the dataset for subsequent periods. For each h, the complete data-cleaning procedure is reapplied, including the removal of singletons and retaining only the largest connected set. All models include worker fixed effects but the figure displays results from three specific models with additional fixed effects. In red, the model includes fixed effects for year, firm, and interaction of collective agreement with occupation. In green, the model incorporates fixed effects for establishment-by-job-title, establishment-by-year, and job-title-by-year combinations. In yellow, the model uses interactions of collective agreement, occupation, year, and establishment as fixed effects, which aligns with our definition of a team.

will be further explored in the subsequent sections.

4.2 Tailored effects: peer quality gaps and worker characteristics

When analyzing a network and considering how observed characteristics influence group formation, empirical research suggests that agents group based on shared attributes. This phenomenon, known as *network homophily*, can generate segmented networks (Jackson, 2021). As we view employer-employee data as a network, attributes as demographic characteristics may not only influence worker allocation but also generate similar responses. Commonly, peer effects research assumes homogeneous social interactions, overlooking how the channel might operate differently across various groups (Comola et al., 2022). This assumption fails to capture the true nature of jobs, which are typically collaborative. Not only productivity depends on coworkers (Mas and Moretti, 2009), but also wages (Cornelissen et al., 2017; Battisti, 2017), and even differences in wages inside the team might affect worker satisfaction (Card et al., 2012). We believe that social pressure may operate differently across various populations, leading to variations in the peer quality gap effect across workers. To disentangle the heterogeneous effects of the peer quality gap, this section applies the methods presented in sections 3.2 and 3.3.

This section aims at disentangling the underlying drivers of peer effects. While linear methods typically analyze heterogeneous effects using sub-samples of the population, in this case, it is crucial to conduct such analyses on a complete sample, because any sampling of the largest connected set inevitably changes the network structure. Our analysis is divided into four parts: *(i)* First, we examine demographic characteristics, focusing on age, gender, and location. *(ii)* Afterwards, we explore the worker's role within the team, specifically the hierarchical position. Differences in effects based on worker positions might impact organizational dynamics and productivity. We begin by focusing on the hierarchical roles of managers and subordinates within teams, followed by an analysis of job content and responses based on worker skill levels. Then we compare workers based on their likelihood of leaving the team, contrasting 'established' and 'vulnerable' workers. *(iii)* Right after we analyze responses based on worker mobility. We group workers based on their transition status within the team, identifying differences in responses to the peer quality gap between incumbents, workers hired from a firm, and those coming from unemployment. *(iv)* Finally in the next section, we group workers based on their unobserved worker quality to determine the responses of *'better'* and *'worse'* workers. This extended analysis enables us to identify the potential drivers of the peer quality gap and its implications for workers and firms.

4.2.1 Demographic characteristics.

Our data source contains detailed information on the worker, enabling a complete demographic characterization. To ensure that our variable choices do not overlap with the definition of the peer group used in the paper, we carefully selected these attributes. In this section, we explore how the peer gap quality varies across age, gender, and work location.

The role age, gender, and experience. Coworkers may not only affect outcomes but also our attitudes within the team, potentially creating heterogeneous effects in our peer quality gap. Examples of this can be found in education, where researchers have observed that peer effects across gender differentiate not only academic outputs but also attitudes and socio-emotional skills (Gong et al., 2021). Gender also appears to be a major determinant influencing occupational choices and career paths. These changes might be attributed to team dynamics, such as competitiveness (Buser et al., 2014), or the structure of the network (Beugnot et al., 2019). We anticipate observing different gender responses in the workplace. While evidence remains limited, Messina et al. (2023) recently demonstrated that gender norms might result in different peer effects responses between and within genders. The authors also point out that peer pressure is the primary driver behind gender differences in peer effects.

Table 3: Heterogeneous Peer Gap Effects by Age and Gender

	Gender		
Age bins	Male	Female	Diff.
18-24	0.542***	0.445***	0.097***
	(0.0014)	(0.0012)	(0.0018)
25-34	0.443***	0.382***	0.06***
	(0.0008)	(0.0008)	(0.0012)
35-49	0.344***	0.299***	0.045***
	(0.0008)	(0.0008)	(0.0012)
50+	0.322***	0.306***	0.017***
	(0.0009)	(0.0009)	(0.0013)

Source: Quadros de Pessoal, MTSSS, $1994-21$.
Note: This table shows the results from running the heterogeneity model with an interaction between age and gender. The column "Male" reports the $\hat{\eta}$ coefficients for males by each age bin; and the column "Female" reports the same for women. Finally, the column "Diff" reports numbers for . The controls are age, tenure and their respective quadratic terms. Clustered (peer group) standard-errors in parentheses. Significance Codes: ***: 0.01, **: 0.05, *: 0.1

One of the reasons to focus on gender is the substantial increase in female labor participation over the past decades. As evidenced by (CITE), females are known to perform better academically and pursue more years of education, which has influenced the overall workforce structure. Figure B2 illustrates the evolution of the workforce structure over the last three decades. While the figure reveals large differences in workforce composition by gender, this change has been gradual. Despite a general aging of the workforce, females have experienced substantial increases in labor force participation compared to men, especially in older age groups. This is primarily due to the historically low absolute numbers of females in the older age range of the labor force at the beginning of the period analyzed. Overall, the share of younger workers has diminished, a result of demographic transition and the high unemployment rate for young workers in Portugal. However, this decrease has been sharper for females. These findings call for a joint analysis of both gender and age.

To evaluate the presence of heterogeneous effects of the peer quality gap by age and gender, we estimate a single non-linear model. Table 3 presents the resulting coefficients of the estimation, which show distinct significant effects for each age-gender group. Females are less sensitive to the peer gap in all age categories, and the sensitivity for young workers is higher than for older workers. Moreover, the responsiveness difference between genders diminishes with age. The latter could indicate that experience might also play a differential role between males and females. The effect of a 1 s.d. increase in the quality gap on wages is 28.3 % for male young workers, and decreases to 12.6 % for the eldest group. For males, the effect is $54.98\ \%$ lower if we compare the sensitivity of old to young workers. While the coefficient for females diminishes less than for men ($-55.6\ \%$). The magnitude of the peer quality gap coefficient implies a wage change ranging from 14.4 % for the youngest female, to 8.9 % for the oldest females in the market. Thus, compared to men, young females are 49.2 % less responsive to peer quality, but as they grow older the gap decreases to 29.4 %. These findings suggest that women may be less responsive to peer effects, which could have consequences for their job-to-job mobility. As Del Bono and Vuri (2011) documented using social security data for Italy, lower worker mobility for females can harm their wage profile and contribute to a widening gender wage gap.

Table 4: Standard Deviations and Marginal Effects by Gender and Age

Age	Gender	Population SD $sd(\alpha_i)$ (1)	Section SD $sd(\bar{\alpha}_{-i} - \alpha_i)$ (2)	Marginal Effect Ratio (3)
All	All	0.459	0.339	0.219
Gender-Age Model				
18-25	Female	0.305	0.302	0.403
18-25	Male	0.407	0.460	0.563
25-34	Female	0.390	0.299	0.254
25-34	Male	0.463	0.396	0.325
35-49	Female	0.423	0.269	0.159
35-49	Male	0.507	0.356	0.197
50+	Female	0.429	0.278	0.166
50+	Male	0.542	0.368	0.175

Source: Quadros de Pessoal, MTSSS, $1994-21$.
Note: This table shows standard deviations and marginal effects calculated using the simple gap model and the model with gender and age response heterogeneity. Column 1 contains the standard deviation of worker fixed effects and column 2 the standard deviation of the average peer quality, both calculated separately by age and gender. "Marginal Effect Ratio" is a ratio that divides the marginal effect on wages from a one standard deviation increase in peer quality by the marginal effect of a 1 standard deviation increase in own quality.

To assess the differentiated gender effect on mobility, we investigate how the quality gap changes with labor market experience. (add to this what is in the intro) Figure 3 shows that the negative quality gap decreases with experience for both genders. When we check the evolution of peer quality since labor market entry, we observe that both males and females increase their peers' quality throughout their careers. However, the increase in peer quality is bigger for males than for females. This has important implications for the gender wage gap, particularly in wage profiles, as evidenced by recent literature (Manning and Swaffield, 2008). The predicted wage rises 17 % more for men than for women as their careers progress. This gap can be explained by differences in tenure gains, due to time out of the labour force, and gains from job mobility Loprest (1992). More specifically, it has been argued that men moving to better firms is the main explanation behind higher gains from job mobility for males (Del Bono and Vuri, 2011; Casarico and Lattanzio, 2024). Although we find that firm switches do explain part of that wage growth, more than 50 % of the gap is explained by a higher increase in team quality for men, not firm quality.

Sensitivity to geographical factors. To evaluate whether peer gap effects vary geographically, we contrast the working population of Lisbon with workers employed elsewhere. This serves as a placebo test for our estimator. While evidence suggests that peer effects can vary significantly depending on geographical areas, we did not find any cultural or institutional differences between the two groups. Previous research has focus on the role of the workers' network across neighborhoods, finding small or null effects (Eliason et al., 2023). Furthermore, our analysis excludes agricultural and construction work, sectors that might exhibit differences in labor institutional arrangements. This excludes a significant portion of rural employment. The rural-urban divide is the only potential source of different effects, but given that our sample predominantly covers urban work, we anticipate homogeneous effects. As shown in Column (5) of Table 5, the estimates are indeed very similar both among themselves and to the estimates of the base regression presented in Table A3. This suggests that geographical factors do not influence the peer gap effects in our sample.

4.2.2 Sensitivity to the team structure.

Work structure and organizational factors can influence how individuals respond to their peers, as the quality peer gap effects depend on the structure of the network itself. Different team structures may lead to varied responses from workers. Consider, for example, a peer group of 10 workers: the degree of peer pressure differs significantly when half of the peer group are supervisors compared to when only one is. We explore the heterogeneous effects of the peer quality gap on workers, depending on their *'status'* within the team. We examine three distinct categories that might exhibit different responses: (i) First, we study if the worker's skill level moderates the effect of the peer quality gap. (ii) We analyze whether there are variations in the impact of the peer quality gap between managers and subordinates. (iii) We investigate whether the peer quality gap affects vulnerable workers (e.g., young, unskilled, recent hires) differently compared to established workers (e.g., old, experienced, high-skilled employees).

Skill level and the peer quality gap Peer pressure is likely the primary driver of the peer quality gap; in this case an individual's peers' effect may vary depending on the nature and difficulty of the tasks they are required to perform. To study how the peer quality gap effect varies across skills, we make use of the ISCO-08 skill classification which associates each occupation with one of three skill levels: high, medium, or low. This skill proxy overlaps with the peer group definition, so comparing skills effectively means comparing between different teams.

Column (6) in Table 5 presents the estimates for the different skill levels. We observe that the coefficients are larger for high-skill workers and decrease for lower skill levels. However, when we compute the effect on wages, the effect becomes quite similar. The wage effect from a one standard deviation increase in the wage quality gap implies a $14.2\ \%$ increase in wages for high-skill workers, $13\ \%$ for medium-skill workers, and $13.5\ \%$ for low-skill workers. These results contrast with recent findings on peer effect estimates in

Table 5: Heterogeneous Peer Gap Effects on Wages

Dependent Variable:	Log of hourly real wage					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Var. interact with $\hat{\bar{\alpha}}_{-i} - \hat{\alpha}_i$						
All	0.355^{***}					
	(0.001)					
Safe		0.353^{***}				
		(0.001)				
Vulnerable		0.403^{***}				
		(0.001)				
Subordinate			0.419^{***}			
			(0.001)			
Supervisor			0.189^{***}			
			(0.001)			
Incumbent				0.338^{***}		
				(0.001)		
Switched jobs				0.358^{***}		
				(0.002)		
Hired from unemp.				0.445^{***}		
				(0.001)		
Other				0.400^{***}		
				(0.001)		
Non-Lisbon					0.358^{***}	
					(0.001)	
Lisbon					0.356^{***}	
					(0.001)	
High-skilled						0.411^{***}
						(0.001)
Medium-skilled						0.379^{***}
						(0.001)
Low-skilled						0.279^{***}
						(0.001)
controls	✓	✓	✓	✓	✓	✓
Fixed-effects						
worker	✓	✓	✓	✓	✓	✓
firm	✓	✓	✓	✓	✓	✓
job title	✓	✓	✓	✓	✓	✓
year	✓	✓	✓	✓	✓	✓
Fit statistics						
Observations	24,046,313	24,046,313	24,046,313	24,046,313	24,046,313	24,046,313
R^2	0.90598	0.90609	0.90633	0.90633	0.90600	0.90634
Within R^2	0.62525	0.62567	0.62662	0.62662	0.62533	0.62668

Source: Quadros de Pessoal, MTSSS, $1986 - 21$.

Note: This table only reports the estimated $\hat{\eta}$ coefficients and fit statistics for four models. Model (1) is a simple estimation, without response heterogeneity, following equation **??**. Models (2), (3), and (4) have response heterogeneity. In Model (1) it relates to the job safety of a worker, calculated by the predicted probability of separating from the current firm. We estimated the probability using a *logit* model with wage, wage squared, [fill with more]. In Model (2) the heterogeneity relates to the worker's role in the firm. We classify a worker as a manager if he's either classified as a manager in its occupation or has a hierarchical position that implied a managing role. In Model (2) the heterogeneity relates to the most recent flux of the worker. An incumbent is a worker that stayed in the same peer group. Switching jobs consists of a separation to another firm. In this case, this is a robust job switch a worker must have been employed for at least 3 years at the firm from which he separated from; lastly, a hiring from unemployed is a worker a not at that firm in the previous year and of a worker that has been *out of sample* for at least three years. The controls are age, tenure and their respective quadratic terms. Clustered (peer group) standard-errors in parentheses. Significance Codes: ***: 0.01, **: 0.05, *: 0.1. TO ADD: description of model 5 and 6.

the Brazilian labor market (Messina et al., 2023), which observe a stronger effect for low-skill jobs. This discrepancy might be attributed to differences in our skill and peer group definitions[15]. Our skill categorization considers the type of tasks and skills required to perform a job. When a job has a high cognitive skill component, it is often complemented by social skills. This can lead to reduced coordination costs for these jobs, increasing job specialization and collaboration (Deming, 2017). These features are characteristic of high-skill jobs in our occupational classification. Our coefficients align with evidence in educational economics regarding stronger peer effects for soft skills (Zárate, 2023). Other disciplines, such as sociology, consider that occupation is tied to class structure. This perspective can be valuable in this context as it represents common values, attitudes, shared beliefs, and social norms that may explain these differences.

By hierarchy inside the peer group. Our interest in studying the hierarchical structure of the peer group stems from the possibility that peer pressure operates also through this channel. In lab experiments, the fact of being monitored doing a task have effects on productivity (Georganas et al., 2015), and the strength of monitoring is related to peer pressure (Battiston et al., 2023). As the management layers increase, or the team structure privilege the supervisor roles, subordinate workers are subject to higher peer pressure. Thus, hierarchy might impose power dynamics, influencing who feels more pressure within a team. Additionally, the structure of a team impacts its productivity and affects how team formation occurs. Given our focus on peer pressure, we anticipate different reactions from supervisors and subordinates. Our data enables us to identify hierarchical layers within peer groups and classify workers based on this structure[16]. Although our definition of peer groups accounts for occupation, we don't use the occupational title. Instead, we divide workers into two layers within the peer group: those who coordinate or supervise and are more independent setting the task at work ('supervisors') and those who comply and are less independent ('subordinates'). Column (3) in Table 5 presents the estimated coefficients. The results align with a peer pressure mechanism, as managers are less sensitive to the peer quality gap. While a one standard deviation increase in the peer quality gap increases wages by $8\,\%$ for managers, it does increases $15.6\,\%$ for subordinates.

Vulnerable workers. Structural parameters of the economy can influence how different groups of workers behave. For example, recent evidence suggests that separations are more sensitive to wages among young, less educated workers (Autor et al., 2023). Building on this idea, we aim to evaluate whether there are differences in responses to the peer quality gap based on how workers compare to each other within their team integration. We focus on comparing well-established workers (low probability of separation) with vulnerable workers (high probability of separation)[17]. Frakes and Wasserman (2021), using information on

[15] In Messina et al. (2023) paper skills are defined using the postgraduate share, rather than considering the nature of the job. There are other differences in both approaches. Our definition of peer groups is more granular as the peer group is defined at the job-title level, instead then at the firm level.

[16] For the construction of the organizational layers we follow Caliendo et al. (2020). The identification of the parameter relies on both the roles between and within the team. While some teams are composed of subordinates and supervisors, we allow compare between groups that are not mixed.

[17] To identify vulnerable workers, we use a logit model to predict the probability of job separation. We use the observed worker separations (whether a worker leaves the job) as the dependent variable. The explanatory variables include worker characteristics such as age, tenure, gender, and wage. For continuous variables, we also include their squares to capture non-linear effects. Additionally, we incorporate log-firm size, sector fixed effects, and occupation fixed effects. Based on the estimated model, we calculate the predicted probability of

examiners at the US Patent and Trademark Office, find that peer effects are stronger at the beginning of careers and decay with experience. The authors attribute this to a signaling mechanism, suggesting that individuals may be less concerned about stigma later in their careers. Consistent with this evidence, we expect peer pressure to influence these groups differently. Apart from having higher tenure and experience, which disclose information about their type, established workers may have consolidated stronger networks within the firm and peer group, potentially reducing their exposure to peer pressure. This could make them less sensitive to the quality peer gap. The coefficients presented in Column (2) of Table 5 align with this expectation, as the coefficient for vulnerable workers is larger than for established workers. In terms of wages, a one standard deviation increase in the quality peer gap implies a wage increase of only $12.5\ \%$ for established workers, compared to a $18\ \%$ increase for vulnerable workers. As a lower effect translates into a lower opportunity cost when the worker remains in the same job, a larger peer quality gap effect implies a higher sensitivity to wages, which can lead to greater job mobility.

4.2.3 Worker mobility and the peer quality gap.

The evidence presented so far highlights a link between the quality peer gap and worker mobility. The results also demonstrate that heterogeneity extends beyond homophily. The heterogeneity does not solely stem from observable characteristics or how workers behave similarly, but also varies based on worker status and situation within the team. To further investigate its heterogeneity in the labor market, we group workers based on their mobility status.

Effect by worker transition status. Given our hypothesis that peer pressure is the primary channel through which the quality peer gap operates, we anticipate that a worker's status within a team will influence their response to the peer quality gap. Drawing a parallel with the experience and signaling argument presented by Frakes and Wasserman (2021), newly hired workers, depending on their origin, may provide different signals to the firm and their peer group. A poached worker, is perceived as competent, at least enough to be in a firm. In contrast, an entry from unemployment carries less information about their quality, potentially leading to higher pressure from colleagues to demonstrate their ability. For incumbent workers, peer pressure may be lower as their peers are already familiar with their quality. The results align well with these explanations. Column (6) of Table 5 presents the coefficients of the peer quality gap estimation for these different groups. A one standard deviation increase in the peer quality gap increases wages by 11.7% for peer group incumbents, and 12.7% for poached workers. The effect is stronger for hires from unemployment (21.1%). These findings strongly support the notion that peer pressure is the underlying channel through which the quality peer gap operates. Furthermore, these results add to the evidence that links peer quality gaps and worker mobility.

4.3 Better workers, worse workers, and team composition.

The previous sections explored how peer effects vary based on observable characteristics and worker status (worker hierarchy, and mobility status within the team). Understanding how different types of workers move can ultimately affect worker allocation to firms,

separation. Workers in the top decile of this predicted probability distribution are then classified as vulnerable.

influencing team formation processes and the final team structure. To properly study the effect of the peer quality gap, it is not enough to focus on those variables. We must consider whether the effects differ when clustering by the peer quality gap itself, as this is the variable where firms can optimize firm formation. This is challenging as existing methods are limited in their ability to estimate such effects. One of the novel methodological contributions of this paper provides a method to estimate heterogeneous treatment effects of the peer quality gap. That is, to group by the unobserved variable of interest. In simple terms, our analysis tries to answer a simple question: can "bad apples" spoil "good apples"? and if so, what is the consequence for team composition? In this application, we divide workers into two sets: "better workers" whose quality exceeds the average peer quality and "worse workers" whose quality is below the average of their colleagues.

Analyzing this heterogeneity sheds light on two aspects: *(i)* How peer pressure (as suggested by previous evidence) influences worker behavior; *(ii)* From the firm's perspective, what is the role of heterogeneous quality in team formation? Considering the direct link between wages and productivity, firms might optimize team composition to benefit from better workers and maximize total effort. This argument aligns with Mas and Moretti (2009), which calculated how different team compositions could improve productivity.

How the peer gap quality is related to team formation? The peer quality gap is directly linked to worker quality variance within the team. Consider first, if peer gap effects are symmetric firms would compensate for the losses of a bad hire with a better worker. In this scenario, the firm strategy is purely a reaction, which results in worker quality dispersion inside the peer group. Dispersion across teams is merely due to luck, leading to random worker allocation. Consequently, when considering the overall labor market, we would find no positive assortativity between workers and firms. However, evidence suggests that peer effects vary across the distribution of worker productivity (Ichino and Falk, 2005). This raises two other possible scenarios. When peer quality gap effects are asymmetric, and the worse workers are more sensitive to its peers, a worse worker could benefit in terms of wages from the presence of better workers, while better workers may not be significantly harmed in terms of wages. If wages and firm productivity are directly linked, increasing the variance within the peer group could enhance productivity at a lower cost. Therefore, firms would be incentivized to maximize quality dispersion within each peer group. Mas and Moretti (2009) documented a similar scenario, where changing the composition of a cashier team in a retail store and increasing its variance reduced costs without any loss in productivity. Ichino and Falk (2005) found similar results using observed output in a controlled field experiment, demonstrating that low-productivity workers benefit more from their peers than high-productivity workers. Lastly, consider a situation where better workers are more sensitive to peer quality gaps. Better workers are severely penalized, negatively affecting their wages and productivity more than worse workers are benefited. This is costly for both the worker and the peer group's productivity. Consequently, firms have an incentive to minimize peer group quality dispersion.

To investigate how peer effects influence worker allocation across peer groups, we estimate a model with heterogeneous effects for better and worse workers in jobs with varying degrees of specialization. Tables A4 presents the results of estimating the asymmetric quality gap effect for better and worse workers at different skill levels. In the table, better workers serve as the base reference in each skill group. To ease the interpretation, Table 6 presents the standard deviations, the instantaneous effect, the log-wage effects and the wage effects for each group. All effects reported are compared to an *"unaffected worker"*, a worker in a team where all members are identical in terms of worker quality, so the peer quality gap is null,

$(\bar{\alpha}_{-i} - \alpha_i) = 0$.

Table 6: Standard Deviations and Marginal Effects by Skill Level

	Population SD $sd(\alpha_i)$ (1)	Section SD $sd(\bar{\alpha}_{-i} - \alpha_i)$ (2)	Marginal effect $\hat{\eta}_d/2 + \hat{\eta}_b$ (3)	-1SD effect on log-wage (4)	+1SD effect on log-wage (5)	-1SD +1SD effect on wages (6)
High-skilled	0.357	0.324	0.445	-0.121	0.117	-0.344
Medium-skilled	0.405	0.322	0.433	-0.116	0.115	-0.060
Low-skilled	0.564	0.453	0.391	-0.150	0.149	-0.094

Source: Quadros de Pessoal, MTSSS, $1994 - 21$.
Note: This table shows standard deviations (SD) and marginal effects calculated using model 5. Column 1 contains the standard deviation of worker FEs and column 2 the standard deviation of the average peer quality, both calculated separately by skill level. "Marginal Effect" is the instantaneous change in log-wages from a change in average peer quality calculated for a worker with a peer quality gap of 0. Also assuming an initial peer quality gap of 0, columns 4 and 5 calculate the marginal effect in log-wages from a one standard deviation increase or decrease in peer quality. Finally, column 6 is the sum of column 4 and 5, transformed to reflect the difference in wage variation between a SD increase and a SD decrease in peer quality gap ($(e^{diff} - 1) \times 100\%$).

Overall, the effects of the η coefficient are larger for "better" workers compared to "worse" workers, and the effect increases with the skill level. For low-skill jobs, a one standard deviation change in the peer quality gap has an almost symmetrical effect on log-wages. For the unaffected worker, a decrease of one standard deviation in peer quality implies a decrease in the peer quality gap that negatively affects wages (-0.150 log-points). Similarly, an increase in peer quality results in an increase in wages of almost equivalent absolute value (+0.149 log-points). If the changes are consecutive, it translates to a small wage decrease of 0.09%. Similar effects can be found for medium-skill jobs, with the effects being very similar in magnitude for the unaffected worker. Because the peer quality gap effects are stronger for better workers, a consecutive increase and decrease of one standard deviation in peer quality would result in a wage loss (-0.06%). When considering high-skill jobs, the effect becomes asymmetric. For the unaffected worker, a decrease in peer quality of one standard deviation has a stronger negative effect on log-wages (-0.121 log-points) than an increase ($+0.117$ log-points). This effect predominates when the effect is consecutive, ultimately shrinking largely the starting wage (-0.34%).

To put these numbers into perspective, consider a firm with four workers employed in high-skill occupations and divided into two teams. The first team comprises two high performers, each 1 standard deviation above the median worker fixed effect, while the second team consists of two low performers, each 1 standard deviation below the median. If the firm reorganizes by having one worker from each team switch places, resulting in each team having one good and one bad worker, average wages would decrease by 1.6%. This reduction is entirely driven by the change in allocation, since the workers themselves remain the same. Importantly, for low-skilled workers, this allocation impact would be more than three times smaller.

For low and medium-skill jobs, worker quality dispersion can be large, as peer effects compensate for each other on both sides of the peer quality gap. However, for high-skill jobs, the sensitivity of better workers to the peer quality gap is larger than that of worse workers. These results imply that having a worse worker is highly penalized compared to the benefits derived from better workers. This suggests that firms have no incentive to have a large peer group quality dispersion, as "bad apples" spoil "good apples". The results suggest that increasing the level of specialization, we expect to observe a lower variance in worker quality. The peer quality gap sheds light on how teams are composed and why good teams often consist of good workers. While positive assortativity is a well-established finding in labor literature (Abowd et al., 1999), and peer effects literature (Battisti, 2017), our results provide an alternative explanation for the underlying channel and how it operates.

4.4 Quality induced mobility

Until this point, this paper has provided evidence on the importance of the peer quality gap. We showed the existence of heterogeneous effects across observable and unobservable characteristics and presented how these effects can create different responses for workers, induced by peer pressure, and for firms as heterogeneous peer effects within the peer group influence team formation. This ultimately affects the allocation of talent and the workforce structure. Such heterogeneous effects serve as evidence that points towards a link between mobility and the peer quality gap. This section investigates how the quality gap influences worker mobility choices. Our aim is to quantitatively assess the importance of different factors that can explain worker mobility, with the peer quality gap being one of the factors analyzed. Understanding worker mobility is relevant as it sheds light on several phenomena: worker allocation, sorting, and earning dynamics. Including the peer quality gap opens new avenues for understanding how these concepts are interrelated.

To understand what factors can induce mobility, we must consider that workers typically change job when it offers better opportunities than their current position. It is challenging to take directly to the data as the decision generally is not based on a unique dimension. Workers may move not only for higher wages but also for the possibility of future wage growth in a more productive firm (Postel–Vinay and Robin, 2002), to use particular skills acquired and developed during their career (Gathmann and Schönberg, 2010; Cortes and Gallipoli, 2018)), or to seek non-compensation-related amenities (Sorkin, 2018). While earlier studies on labor mobility primarily focused on its relation to the economic cycle, we now have a better understanding of mobility across occupations and sectors. Groes et al. (2015) found that occupational worker mobility is linked to ability, with the probability of changing occupation increasing in both tails, and the direction of movement often aligning with worker ability (better workers move to better occupations and vice versa). Such occupational changes are reflected in how workers transition between industries, as the rates of worker mobility change with the occupational structure (Hervé, 2023). Mobility can be influenced by tenure, age, ability, and the specific skills acquired in the labor market. Using pre-hiring data on multiple cognitive and social skills, recent evidence by Fredriksson et al. (2018) links match quality to job mobility. This aligns with our research, although our quality gap is based on a constant unobserved characteristic, while Fredriksson et al. use a multi-dimensional approach.

To further highlight the importance of the peer quality gap as a determinant wage growth, Panel (a) of Figure 3 presents the determinants of wage growth. Workers increase their wages by $66\ \%$ over their careers.[18] The figure shows that factors associated with worker mobility explain a third of wage progression, and changing colleagues accounts for two-fifths of that growth. Surprisingly, this effect is more than double than that attributed to firm mobility, suggesting that the quality of coworkers, rather than the specific firm, may be a more significant factor for wage growth. Panel (b) of Figure 3 highlights the importance of the peer quality gap in explaining the gender pay gap. The largest contributor to the gender pay growth difference is peer quality (31%). Overall, these descriptive statistics suggest how

[18] In line with Manning and Swaffield (2008), we find that males and females have different wage profiles. In our sample, the final average male wage increases by $89\ \%$ compared to entry, while for females, it increases by only 67%.

the peer quality gap operates, leading to differential effects on wages and mobility across genders, which ultimately contributes to the gender pay gap. Motivated by these patterns, we next investigate whether coworkers can directly influence the mobility decision.

To disentangle the effects of various factors on worker mobility, we follow a similar approach to Groes et al. (2015). Our goal is to calculate the conditional effects of each variable of interest on the mobility likelihood. We focus on age, tenure, firm pay premium, the residualized wage, and the peer quality gap. To capture these effects accurately, we restrict our sample to workers within the firm's job-to-job mobility network, focusing on workers in firms that have experienced a hiring or separation event during the period t. We further restrict our sample to workers below 50 years old to avoid including early retirees and to workers that appear in our sample between 1994-1997 and 2018-2021. Using the results of the previous estimation, we can decompose the wage into two parts. The first part is the sum of the attributes of the worker: its quality. The second part contains the attributes of the job, composed of the residual, the interaction of collective agreement with occupation, the firm pay premium, and the peer gap effect by skill level. We use the job-related attributes as a proxy for job quality, assuming that workers choose jobs based on these factors. To better understand the different types of mobility, we divide worker transitions into four groups: those who leave private employment (*'no job'*), those who move to jobs with a 5% or higher job quality (*'better'*), those who move to jobs with a 5% or lower job quality (*'worse'*), and the remaining cases.

In order to quantitatively assess how different worker and job characteristics affect the likelihood of changing jobs, we employ a Generalized Additive Model – GAM (Hastie and Tibshirani, 1987; Wood et al., 2015). A GAM is an extension of a Generalized Linear Model (GLM), where the linear predictor (the combination of explanatory variables) is replaced by a sum of smooth functions of the predictors. In this case, the outcome is a binary variable, indicating whether a worker moved jobs in period $t+1$. The set of dependent variables x_q, which include variables like age, tenure, residualized wage, firm pay premium, and peer quality gap, are incorporated through smooth functions $f(x)$ to capture nonlinear relationships. We don't use the independent variables directly but instead calculate its percentile on each distribution in its relevant labor market, $\tilde{x}_q$. The relevant labor market is defined as the workers in the same occupation, year, and geographical location using the NUTS2 classification. This insures that the variables are comparable between themselves and occupations. We estimate the following model:

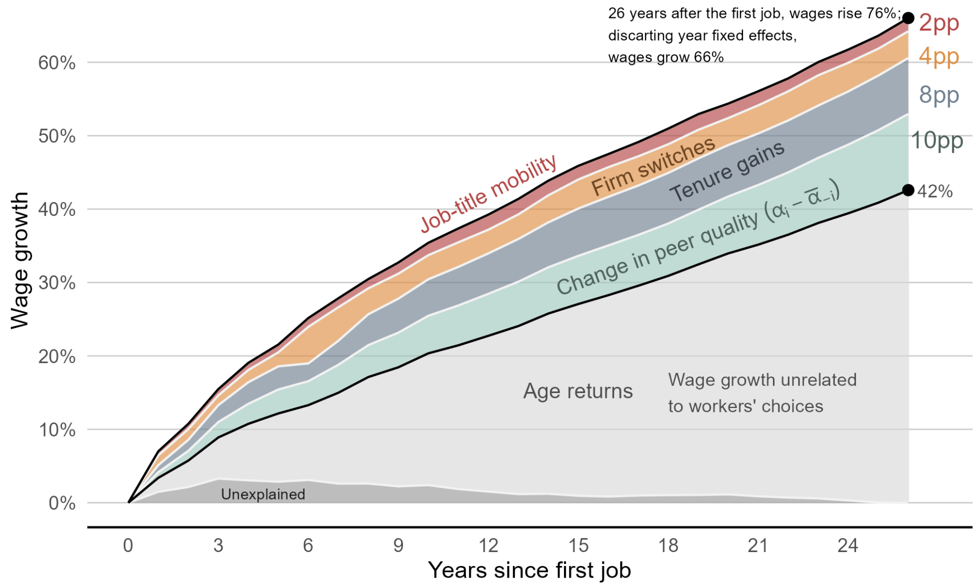
$$\Pr(\textit{separation}_{i,t+1} = 1) = \alpha + \sum^{q} f_q(\tilde{x}_{qit}) + \varepsilon_i. \tag{6}$$

For all the variables, the smoothing function, f, is a thin plate spline, with a single second derivative in its penalty. Model (6) allow us to understand how the likelihood for the different transition types varies across the distribution of each observable, while conditioning on the effects of other variables in the model.

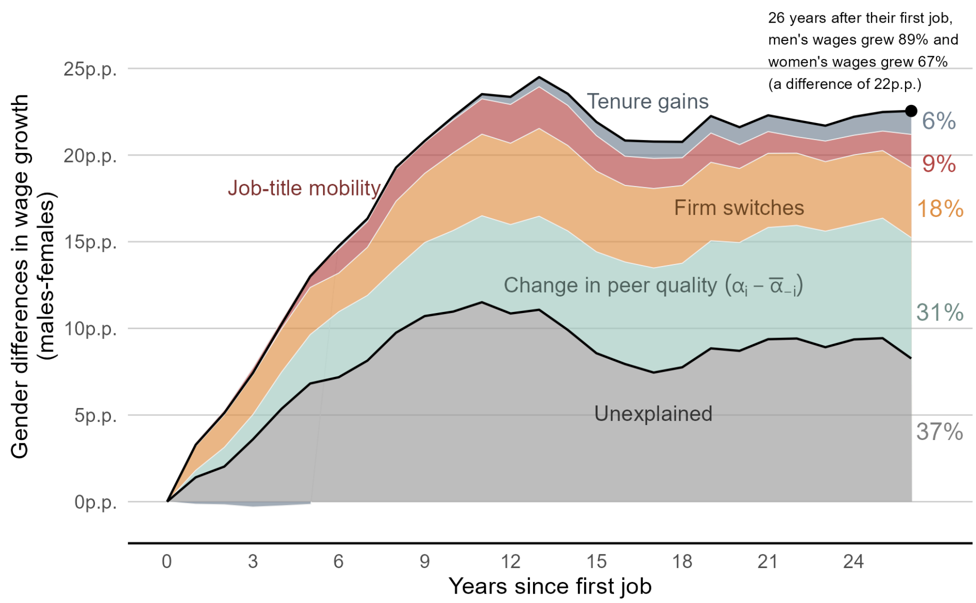
Panels (a) and (b) in Figure 4 present the results of the estimation of equation (6). All the figures describe how the probability of worker mobility change across the distribution of observables. The curves are a non parametric plot of the probability that a worker moves. The curves are conditional to the effect of the rest of covariates. For all the figures, we fix the predicted value at the 50^{th} percentile for all the variables, and let the variable of interest vary across percentiles.

Panel (b) in Figure 4 shows the behavior of the residualized wage, firm pay premium, age, and tenure. For all types of mobility, and for all variables, the probability of leaving private

Figure 3: Importance of Peer Effects in Workers' Careers



(a) Decomposing Wage Growth

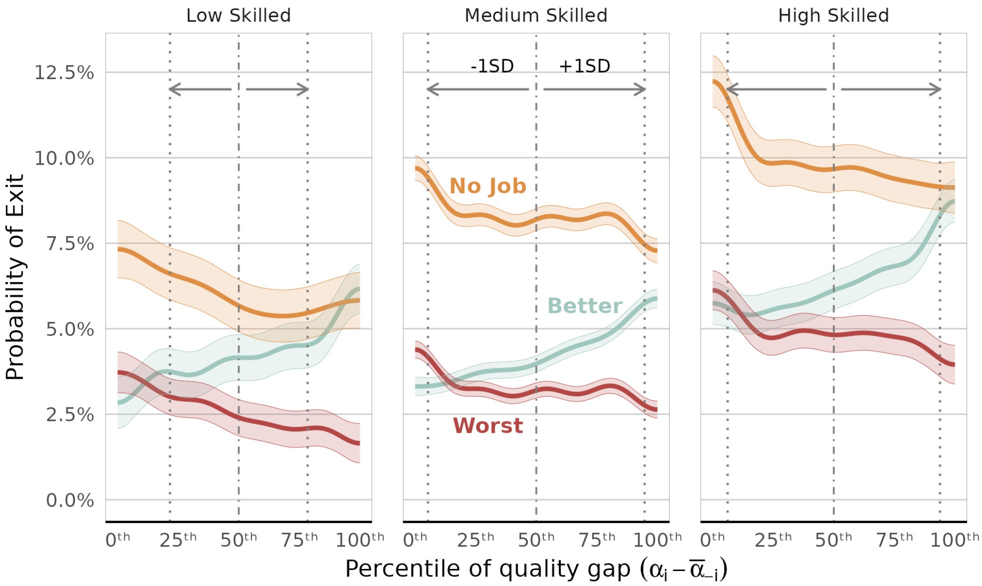


(b) Decomposing the Gender Wage Growth Gap

Source: Quadros de Pessoal, MTSSS, $1994 - 21$.

Note: This figure illustrates wage growth over time (Panel A) and compares wage growth differences between males and females (Panel B), segmented by different wage components. The analysis includes only those workers who were employed between 1994 and 1997 and again between 2018 and 2021. This selection ensures a relatively consistent sample of workers, minimizing significant compositional changes while avoiding overly strict exit conditions that could disproportionately exclude certain groups, particularly women during childbirth. Wage growth is calculated as follows. First, we first estimate model X, which accounts for heterogeneity in response by age and gender. Second, we fit the predicted components from this model. The components are: age, residuals, year fixed effects, tenure (indicating non-mobility, or staying with the same firm), changes in the interaction of collective agreement with occupation, firm switches, and changes in peer group quality (which can relate to the movement of the worker or the movement of colleagues). Third, for each component, we calculate the average value by years of experience and subtract the average value at the start of the worker's career (i.e., when experience is 0). Finally, since wages are expressed in logarithms, we exponentiate the values and subtract 1 to obtain the percentage wage growth. For panel B, Age and year FE account for almost nothing of the difference. Third, for each component, we calculate the average value by years of since the first job and subtract the average value at the start of the worker's career (i.e., when experience is 0). Finally, since wages are expressed in logarithms, we apply the following formula to obtain the percentage points difference: $g_m - g_f = [1 + g_f] \times [e^{\text{sum of components}} - 1]$, to which we then divide with the proportion of each component.

Figure 4: Determinants of Worker Mobility



(a) Quality Gap

(b) Covariates

Source: Quadros de Pessoal, MTSSS, 1994 – 21.

Note: This figure examines the determinants of worker mobility by presenting the predicted values extracted from model 6. Panel (a) illustrates results for the quality gap by skill level, while panel (b) focuses on other covariates. For each variable, percentiles within a local labor market (defined by NUTS 2 region, year, and occupation) are calculated and included as inputs to model 6. Shaded regions in the figure indicate 95% confidence intervals. The predicted values correspond to three distinct dependent variables utilized in model 6. The category "no job," shown in yellow, represents workers who were not recorded in QP during the following year. The category "better," depicted in green, indicates workers who transitioned to higher-paying jobs, while "worse," shown in red, reflects those moving to lower-paying positions. This analysis includes only those workers who were employed between 1994 and 1997 and again between 2018 and 2021. This selection ensures a relatively consistent sample of workers, minimizing significant compositional changes while avoiding overly strict exit conditions that could disproportionately exclude certain groups, particularly women during childbirth.

employment is higher. However, this is tied to our mobility type definition. The top left panel of the figure shows that the probability of moving to a better job sharply decreases as the residualized wage increases in the relevant labor market. The opposite occurs for transitions to worse jobs. Both behaviors can be explained by relative scarcity. It is more difficult to find a higher-paying job when you are already well-paid. In the top right panel shows that the probability of leaving private employment does not differ along the distribution of the residualized wage. The probability of moving to a better job sharply decreases when moving from the first to the second quartile of the firm pay premium distribution. When workers move up in the distribution, there is a smooth decrease. This aligns with firm-worker assortativity, as better firms tend to pay better. The probability of leaving private employment decreases slowly as we move up in the firm quality percentile. Below the 75th percentile of the firm pay premium distribution, workers are more likely to move to better jobs than to worse ones. This pattern reverses in the top quartile of the distribution, where workers are more likely to transition to worse jobs. The bottom left panel describes how the likelihood varies with age. Controlling for other covariates, the effect of age remains constant across the distribution. For all ages, workers are more likely to move to better jobs than to worse ones. The bottom right panel shows how the likelihood of each type of move decreases along the tenure distribution. The decrease is sharp in the first quartile and smooth as we move to the top of the tenure distribution. For all tenure levels, workers are more likely to move to a better job than to a worse one.

Panel (a) in Figure 4 presents how mobility varies along the peer quality gap for low, medium, and high-skill workers. To ease interpretation, we reverse the scale of the peer quality gap: at the bottom of the distribution, we observe workers worse than their peers, and at the top, we observe workers better than their peers. For all skill levels, the probability of changing to a better job increases as the peer quality gap increases, with a stronger effect in the top quartile. Conversely, the likelihood of moving to a worse job increases when moving towards the bottom of the distribution. While the signal is noisy at the bottom of the distribution, making it difficult to disentangle which likelihood dominates, above the first quartile, the trend is clear: the likelihood of moving to a better job is always higher than moving to a worse team. Moreover, this gap increases as we move up in relative quality. For medium and high-skill jobs, the probability of moving to a worse job decreases in both tails of the peer quality gap distribution, while it remains constant in the middle. The probability of leaving private employment decreases as we move up in the distribution but increases with skill level. For medium and high-skill jobs, there is a sharp decrease along the first quartile.

The result presented in previous section, in line with equation (2), implied there is a wage penalty when workers are better than their peers. As relative better workers move to better jobs, and vice-versa, the results presented here show that the peer quality gap not only impact wages, it induces mobility and improve worker sorting.

5 Conclusion

Using matched employer-employee data from Portugal, this paper investigates the impact of the peer quality gap on wages and mobility. By leveraging a novel methodological approach, we calculate the effects of the peer quality gap on different populations, considering both observable and unobservable characteristics. These results serve as evidence on how the peer quality is linked to worker mobility. Our findings shed light on how the peer quality gap can shape team formation and individual career trajectories.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333. (Cited on page(s) 2, 3, 9, 26)

_, Robert H. Creecy, and Francis Kramarz, "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," Technical Report 2002-06, Center for Economic Studies, U.S. Census Bureau March 2002. Publication Title: Longitudinal Employer-Household Dynamics Technical Papers. (Cited on page(s) 5)

Aiyar, Shekhar and Christian Ebeke, "The Impact of Workforce Aging on European Productivity," *IMF Working Papers*, 2016, *16* (238), 1. (Cited on page(s) 7)

Andrews, M. J., L. Gill, T. Schank, and R. Upward, "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2008, *171* (3), 673–697. (Cited on page(s) 4)

Angrist, Joshua D., "The perils of peer effects," *Labour Economics*, October 2014, *30*, 98–108. (Cited on page(s) 7)

Arcidiacono, Peter, Gigi Foster, Natalie Goodpaster, and Josh Kinsler, "Estimating spillovers using panel data, with an application to the classroom," *Quantitative Economics*, 2012, *3* (3), 421–470. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/QE145. (Cited on page(s) 3, 8)

Autor, David, Arindrajit Dube, and Annie McGrew, "The Unexpected Compression: Competition at Work in the Low Wage Labor Market," March 2023. (Cited on page(s) 22)

Battisti, Michele, "High wage workers and high wage peers," *Labour Economics*, June 2017, *46*, 47–63. (Cited on page(s) 1, 2, 3, 8, 13, 17, 26)

Battiston, Diego, Jordi Blanes i Vidal, Tom Kirchmaier, and Katalin Szemeredi, "Peer pressure and manager pressure in organisations," *CEP Discussion Papers*, June 2023. Number: dp1924 Publisher: Centre for Economic Performance, LSE. (Cited on page(s) 22)

Beugnot, Julie, Bernard Fortin, Guy Lacroix, and Marie Claire Villeval, "Gender and peer effects on performance in social networks," *European Economic Review*, April 2019, *113*, 207–224. (Cited on page(s) 18)

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa, "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 2019, *87* (3), 699–739. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA15722. (Cited on page(s) 4)

Bono, Emilia Del and Daniela Vuri, "Job Mobility and the Gender Wage Gap in Italy," *Labour Economics*, January 2011, *18* (1), 130–142. (Cited on page(s) 19, 20)

Brune, Lasse, Eric Chyn, and Jason Kerwin, "Peers and Motivation at Work | Journal of Human Resources," *Journal of Human Resources*, July 2022, *57* (4), 1147–1177. (Cited on page(s) 7)

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek, "Gender, Competitiveness, and Career Choices *," *The Quarterly Journal of Economics*, August 2014, *129* (3), 1409–1447. (Cited on page(s) 18)

Cahuc, Pierre and François Fontaine, "On the Efficiency of Job Search with Social Networks," *Journal of Public Economic Theory*, 2009, *11* (3), 411–439. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9779.2009.01416.x. (Cited on page(s) 3)

Caliendo, Lorenzo, Giordano Mion, Luca David Opromolla, and Esteban Rossi-Hansberg, "Productivity and Organization in Portuguese Firms," *Journal of Political Economy*, November 2020, *128* (11), 4211–4257. Publisher: The University of Chicago Press. (Cited on page(s) 22)

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez, "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction," *American Economic Review*, May 2012, *102* (6), 2981–3003. (Cited on page(s) 17)

_ , Ana Rute Cardoso, and Patrick Kline, "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women," *Quarterly Journal of Economics*, May 2016, *131* (2), 633–686. (Cited on page(s) 4)

_ , Jörg Heining, and Patrick Kline, "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics*, 2013, *128* (3), 967–1015. (Cited on page(s) 4)

Cardoso, Ana Rute, Paulo Guimarães, Pedro Portugal, and Hugo Reis, "Human Capital Spillovers and Returns to Education," Technical Report 2022. (Cited on page(s) 1, 3, 5, 8, 10, 13, 14, 15)

Casarico, Alessandra and Salvatore Lattanzio, "What Firms Do: Gender Inequality in Linked Employer-Employee Data," *Journal of Labor Economics*, April 2024, *42* (2), 325–355. Publisher: The University of Chicago Press. (Cited on page(s) 20)

Comola, Margherita, Rokhaya Dieye, and Bernard Fortin, "Heterogeneous peer effects and gender-based interventions for teenage obesity," *CIRANO Working Papers*, September 2022. Number: 2022s-25 Publisher: CIRANO. (Cited on page(s) 4, 8, 17)

Cornelissen, Thomas, Christian Dustmann, and Uta Schönberg, "Peer Effects in the Workplace," *American Economic Review*, February 2017, *107* (2), 425–456. (Cited on page(s) 1, 3, 5, 7, 8, 13, 14, 17)

Correia, Sergio et al., "A feasible estimator for linear models with multi-way fixed effects," *Preprint at http://scorreia. com/research/hdfe. pdf*, 2016. (Cited on page(s) 10)

Cortes, Guido Matias and Giovanni Gallipoli, "The Costs of Occupational Mobility: An Aggregate Analysis," *Journal of the European Economic Association*, April 2018, *16* (2), 275–315. (Cited on page(s) 3, 26)

Deming, David J., "The growing importance of social skills in the labor market," *Quarterly Journal of Economics*, 2017, *132* (4), 1593–1640. (Cited on page(s) 22)

Eeckhout, Jan, "Sorting in the labor market," *Annual Review of Economics*, 2018, *10*, 1–29. Publisher: Annual Reviews. (Cited on page(s) 4)

Eliason, Marcus, Lena Hensvik, Francis Kramarz, and Oskar Nordström Skans, "Social connections and the sorting of workers to firms," *Journal of Econometrics*, April 2023, *233* (2), 468–506. (Cited on page(s) 4, 20)

Falk, Armin and Andrea Ichino, "Clean Evidence on Peer Effects," *Journal of Labor Economics*, January 2006, *24* (1), 39–57. Publisher: The University of Chicago Press. (Cited on page(s) 1, 3)

Frakes, Michael D. and Melissa F. Wasserman, "Knowledge spillovers, peer effects, and telecommuting: Evidence from the U.S. Patent Office," *Journal of Public Economics*, June 2021, *198*, 104425. (Cited on page(s) 22, 23)

Fredriksson, Peter, Lena Hensvik, and Oskar Nordström Skans, "Mismatch of Talent: Evidence on Match Quality, Entry Wages, and Job Mobility," *American Economic Review*, November 2018, *108* (11), 3303–3338. (Cited on page(s) 26)

Galichon, Alfred, *Optimal transport methods in economics*, Princeton University Press, 2018. (Cited on page(s) 4)

Gathmann, Christina and Uta Schönberg, "How General Is Human Capital? A Task-Based Approach," *Journal of Labor Economics*, 2010, *28* (1), 1–49. Publisher: [The University of Chicago Press, Society of Labor Economists, NORC at the University of Chicago]. (Cited on page(s) 3, 26)

Georganas, Sotiris, Mirco Tonin, and Michael Vlassopoulos, "Peer pressure and productivity: The role of observing and being observed," *Journal of Economic Behavior & Organization*, September 2015, *117*, 223–232. (Cited on page(s) 22)

Gong, Jie, Yi Lu, and Hong Song, "Gender Peer Effects on Students' Academic and Noncognitive Outcomes: Evidence and Mechanisms," *Journal of Human Resources*, July 2021, *56* (3), 686–710. Publisher: University of Wisconsin Press Section: Articles. (Cited on page(s) 17)

Granovetter, Mark, "The Strength of Weak Ties: A Network Theory Revisited," *Sociological Theory*, 1983, *1*, 201–233. Publisher: [American Sociological Association, Wiley, Sage Publications, Inc.]. (Cited on page(s) 3)

Granovetter, Mark S., "The Strength of Weak Ties," *American Journal of Sociology*, 1973, *78* (6), 1360–1380. Publisher: The University of Chicago Press. (Cited on page(s) 3)

Groes, Fane, Philipp Kircher, and Iourii Manovskii, "The U-Shapes of Occupational Mobility," *Review of Economic Studies*, April 2015, *82* (2), 659–692. (Cited on page(s) 3, 26, 27)

Hastie, Trevor and Robert Tibshirani, "Generalized Additive Models: Some Applications," *Journal of the American Statistical Association*, June 1987, *82* (398), 371–386. Publisher: ASA Website _eprint: https://www.tandfonline.com/doi/pdf/10.1080/01621459.1987.10478440. (Cited on page(s) 27)

Hervé, Justine, "Specialists or generalists? Cross-industry mobility and wages," *Labour Economics*, October 2023, *84*, 102391. (Cited on page(s) 26)

Hong, Long and Salvatore Lattanzio, "The Peer Effect on Future Wages in the Workplace," SSRN Scholarly Paper ID 4052587, Social Science Research Network, Rochester, NY March 2022. (Cited on page(s) 1, 8, 13, 15)

Ichino, Andrea and Armin Falk, "Clean evidence on peer effects," *Journal of Labor Economics*, 2005, *24*. (Cited on page(s) 24)

Ichinowski, Casey, Kathryn Shaw, and Giovanna Prennushi, "The effects of human resource management practices on productivity: A study of steel finishing lines," *American Economic Review*, 1997, *87* (3), 291–313. (Cited on page(s) 7)

Jackson, Matthew O., "Inequality's Economic and Social Roots: The Role of Social Networks and Homophily," March 2021. (Cited on page(s) 17)

Jarosch, Gregor, Ezra Oberfield, and Esteban Rossi-Hansberg, "Learning From Coworkers," *Econometrica*, 2021, *89* (2), 647–676. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA16915. (Cited on page(s) 3, 4)

Kandel, Eugene and Edward P. Lazear, "Peer Pressure and Partnerships," *Journal of Political Economy*, 1992, *100* (4), 801–817. Publisher: University of Chicago Press. (Cited on page(s) 1, 7, 8)

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten, "Leave-Out Estimation of Variance Components," *Econometrica*, 2020, *88* (5), 1859–1898. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA16410. (Cited on page(s) 4, 5)

Lazear, Edward P. and Kathryn L. Shaw, "Personnel Economics: The Economist's View of Human Resources," *Journal of Economic Perspectives*, December 2007, *21* (4), 91–114. (Cited on page(s) 1)

Leitao, Martim, Jaime Montana, and Joana Silva, "The Role of Firms in Wage Inequality Dynamics," May 2022. (Cited on page(s) 4)

Lindenlaub, Ilse, "Sorting multidimensional types: Theory and application," *Review of Economic Studies*, 2017, *84* (2), 718–789. (Cited on page(s) 4)

_ **and Fabien Postel-Vinay**, "Multidimensional Sorting under Random Search," *Journal of Political Economy*, December 2023, *131* (12), 3497–3539. Publisher: The University of Chicago Press. (Cited on page(s) 4)

Loprest, Pamela J., "Gender Differences in Wage Growth and Job Mobility," *The American Economic Review*, 1992, *82* (2), 526–532. Publisher: American Economic Association. (Cited on page(s) 20)

Manning, Alan and Joanna Swaffield, "The Gender Gap in Early-Career Wage Growth," *The Economic Journal*, July 2008, *118* (530), 983–1024. (Cited on page(s) 19, 26)

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, July 1993, *60* (3), 531–542. (Cited on page(s) 7)

Mas, Alexandre and Enrico Moretti, "Peers at Work," *The American Economic Review*, 2009, *99* (1), 112–145. Publisher: American Economic Association. (Cited on page(s) 1, 3, 7, 8, 17, 24)

Messina, Julián, Anna Sanz de Galdeano, and Anastasia Terskaya, "Birds of a Feather Earn Together. Gender and Peer Effects at the Workplace," *Gender and Peer Effects at the Workplace (November 10, 2023)*, 2023. (Cited on page(s) 1, 3, 4, 8, 15, 18, 22)

Montgomery, James D., "Job Search and Network Composition: Implications of the Strength-Of-Weak-Ties Hypothesis," *American Sociological Review*, 1992, *57* (5), 586–596. Publisher: [American Sociological Association, Sage Publications, Inc.]. (Cited on page(s) 3)

Postel–Vinay, Fabien and Jean-Marc Robin, "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 2002, *70* (6), 2295–2350. (Cited on page(s) 1, 3, 15, 26)

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter, "Firming up Inequality," *Quarterly Journal of Economics*, February 2019, *134* (1), 1–50. Publisher: Oxford University Press, eprint = https://academic.oup.com/qje/article-pdf/134/1/1/28921941/qjy025.pdf. (Cited on page(s) 4)

Sorkin, Isaac, "Ranking Firms Using Revealed Preference," *Quarterly Journal of Economics*, August 2018, *133* (3), 1331–1393. (Cited on page(s) 26)

Tan, Joanne, "Multidimensional Heterogeneity and Matching in a Frictional Labor Market - An Application to Polarization," January 2023. (Cited on page(s) 4)

Wood, Simon N., Yannig Goude, and Simon Shaw, "Generalized Additive Models for Large Data Sets," *Journal of the Royal Statistical Society Series C: Applied Statistics*, January 2015, *64* (1), 139–155. (Cited on page(s) 27)

Zárate, Román Andrés, "Uncovering Peer Effects in Social and Academic Skills," *American Economic Journal: Applied Economics*, July 2023, *15* (3), 35–79. (Cited on page(s) 22)

A All tables of the paper

A.1 Tables

Table A1: Sample Restrictions on the Original Data

Stage	Observations
Original Data	71585372
Drop residual IRCTs	71524678
From 1994 to 2021	59845180
Normal hours $\geq$ 120	54600421
Drop agriculture, fishing and construction	48926999
Age of the worker 16-64	48498911
Remove missing values	47219225
Drop workers with no remuneration	47218289
Base wage $>$ half of the minimum wage	47207301
Drop workers with sick leave or maternity	47207301
Remove self-employed workers	47153733
Tenure $<$ 600 months	47035882
Drop unclassified in qualification	46636076
Establishment size $<$ 2000	45742304
Drop perfectly classified	45742304
Delete peer group singletons	36931452
Largest connected set	24046313

Source: Quadros de Pessoal, MTSSS, $1986-21$.
Note: This table reports the number of observations in the sample after each sample restriction. We delete years 85-93 due to lower data quality. The largest connected set is the largest group of worker-peer group pairs. We delete worker singletons once before deleting peer group singletons, which leads to the existence of worker singletons in the sample. For computational reasons, we restrict establishment size to a maximum of 2000. The normal hours restriction deletes all part-time employees.

Table A2: Literature on Peer Quality Spillovers in the Workplace

	Setting		Peer Group Definition			Results	
Paper	**Area**	**Years**	**Workplace**	**Occupation**	**F. E.**	**Coeff**	**1 s.d.**
Cardoso et al. (2018)	Portugal	1994 - 2013	Estab	3-digit and CA	F/O + Y	0.5000	5.7%
Cardoso et al. (2022)	Portugal	1994 - 2013	Estab	3-digit and CA	E/O/Y	0.2024	2.1%
Hong and Lattanzio (2022)	Veneto, Italy	1975 - 2001	Firm	3 types only	F/Y + F/O + Y/O	0.2580	4.6%
Cornelissen et al. (2017)	Munich, Germany	1989 - 2005	Estab	3-digit	F/Y + F/O + Y/O	0.0110	0.3%
Battisti (2017)	Veneto, Italy	1982 - 2001	Firm	—	F + Y	0.3580	7.8%

Note: This table reports estimates from all the papers that estimate how peer quality affects wages in the workplace. We exclude estimates of heterogeneous effects. Legend of the letters: F for Firm; E for establishment; O for occupation; Y for year, CA for collective agreement. The 3 types of occupation in Hong and Lattanzio (2022) are blue collar, white collar, and managers. The "Setting" section describes the location and time frame of the sample. The "Peer Group Definition" section reports how peer groups are defined. The "FE" column reports which FE were used, and "/" means that the authors used an interaction between fixed effects. Lastly, the section "Results" reports the coefficient of peer effects along with the effect on wages from a 1 standard deviation increase in average peer quality.

Table A3: Peer Effects on Wages: Classic vs. Quality Gap

Dependent Variable:	Log of hourly real wage							
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
$\hat{\bar{\alpha}}_{-i}$	0.93***		0.56***		0.21***		0.29***	
	(0.0005)		(0.001)		(0.001)		(0.002)	
$\hat{\bar{\alpha}}_{-i} - \hat{\alpha}_i$		0.50***		0.36***		0.18***		0.23***
		(0.0002)		(0.0007)		(0.0009)		(0.001)
controls	✓	✓	✓	✓	✓	✓	✓	✓
Fixed-effects								
worker	✓	✓	✓	✓	✓	✓	✓	✓
year	✓	✓	✓	✓				
firm			✓	✓				
job title			✓	✓				
occupation-year					✓	✓		
establishment-occupation					✓	✓		
establishment-year					✓	✓		
peer group							✓	✓
Fit statistics								
Observations	24,046,313	24,046,313	24,046,313	24,046,313	24,046,313	24,046,313	24,046,313	24,046,313
R^2	0.89	0.89	0.91	0.91	0.94	0.94	0.94	0.94
Within R^2	0.89	0.89	0.63	0.63	0.65	0.65	0.65	0.65

Source: Quadros de Pessoal, MTSSS, $1986 - 21$.
Note: This table shows the coefficients and fit statistics for eight models. The models differ by mode of evaluating the peer effects and by fixed effects. Models 1, 3, 5, and 7 use the classic peer effects measure, while the remaining ones use the negative quality gap (average quality of the peers minus own quality). Column four is shaded because it's our preferred specification. The controls are age, tenure and their respective quadratic terms. Clustered (peer group) standard-errors in parentheses. Significance Codes: ***: 0.01, **: 0.05, *: 0.1

Table A4: Endogenous Peer Gap Effects with Skill Heterogeneity

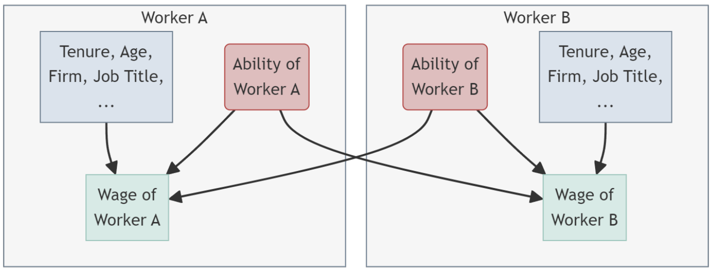
Dependent Variable:	Log of hourly real wage
Model:	(1)
Variables	
$\hat{\hat{\alpha}}_{-i} - \alpha_i \times$ High-skilled $\times$ Worse	-0.0664^{***}
	(0.0033)
$\hat{\hat{\alpha}}_{-i} - \alpha_i \times$ Medium-skilled $\times$ Worse	-0.0117^{***}
	(0.0023)
$\hat{\hat{\alpha}}_{-i} - \alpha_i \times$ Low-skilled $\times$ Worse	-0.0093^{***}
	(0.0025)
$\hat{\hat{\alpha}}_{-i} - \alpha_i \times$ High-skilled	0.4011^{***}
	(0.0020)
$\hat{\hat{\alpha}}_{-i} - \alpha_i \times$ Medium-skilled	0.3654^{***}
	(0.0015)
$\hat{\hat{\alpha}}_{-i} - \alpha_i \times$ Low-skilled	0.3345^{***}
	(0.0017)
age	0.0305^{***}
	(0.0000)
age^2	-0.0002^{***}
	(0.0000)
tenure	0.0009^{***}
	(0.0000)
tenure^2	0.0000^{***}
	(0.0000)
Fixed-effects	
worker	✓
firm	✓
job title	✓
year	✓
Fit statistics	
Observations	24,046,313
R^2	0.90606
Within R^2	0.62555

Source: Quadros de Pessoal, MTSSS, $1986 - 21$.
Note: This table shows the coefficients and fit statistics from model 5. Clustered (peer group) standard-errors in parentheses. Significance Codes: ***: 0.01, **: 0.05, *: 0.1

B All Figures of the paper

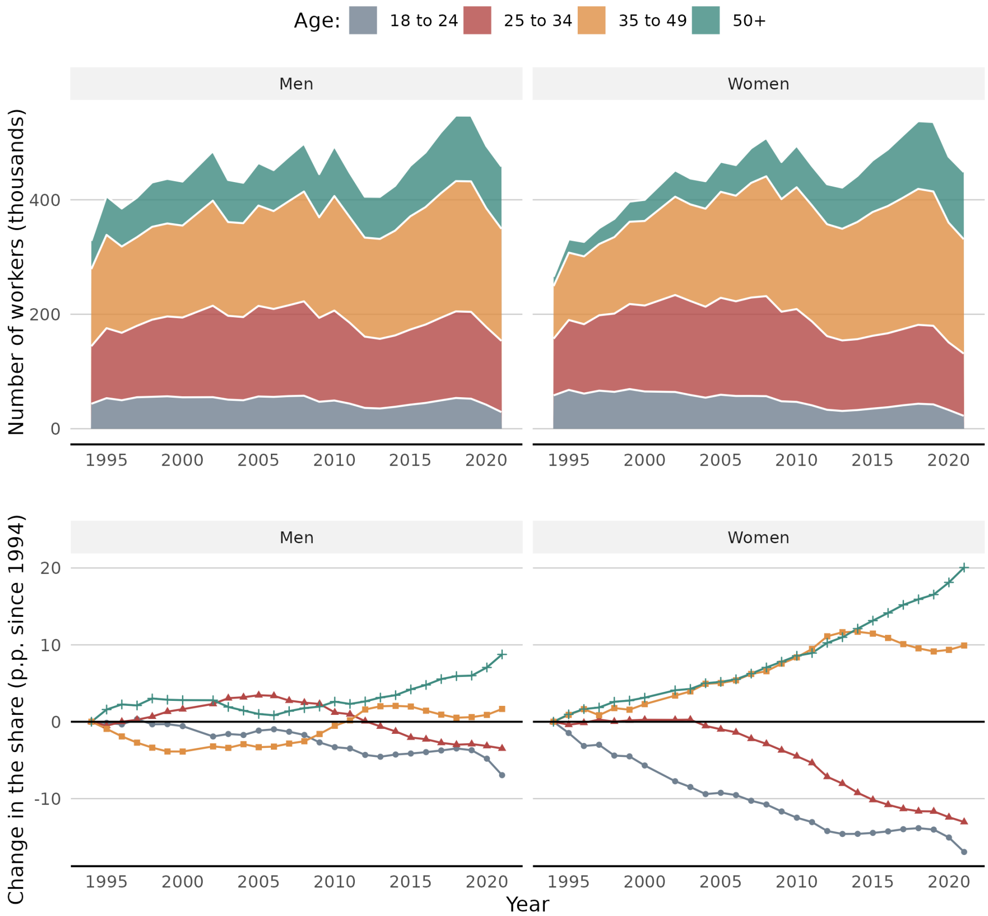
B.1 Figures

Figure B1: Example of Peer Effects in a Two-Worker Team



Note: This diagram illustrates the causal chain for wage setting of workers. In the example, a team consists of two workers, A and B. The dependent variable (in green) represents the wages, which is the outcome we aim to explain (dependent variable). The observable independent variables are shown in blue, while the key unobservable variable, the workers' ability level, is highlighted in red. The ability level not only influences each worker's wage individually but also affects the wages of all their colleagues.

Figure B2: Workforce Evolution by Age and Gender



Source: Quadros de Pessoal, MTSSS, 1986 – 21.
Note: This figure shows the age composition of the Portuguese workforce over the years and by gender. The left-hand side panel represents males and the right-hand side represents women. In blue we represent any worker between 18 and 14 years old; workers between 25 and 34 years old are in red; between 35 and 49 are in yellow and in green we show all the workers with 50 or more years. The upper panel shows the absolute number of workers in each gender/age-bin group, in thousands. The bottom panel shows the change in percentage points of each bin comparing with 1994. For example, if a bin of 20% of women became 30%, the bottom panel would show a 10 p.p. increase.