I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16721

Birds of a Feather Earn Together. Gender and Peer Effects at the Workplace

Julián Messina
Anna Sanz-de-Galdeano
Anastasia Terskaya

JANUARY 2024

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16721

Birds of a Feather Earn Together. Gender and Peer Effects at the Workplace

Julián Messina
University of Alicante and IZA

Anna Sanz-de-Galdeano
University of Alicante and IZA

Anastasia Terskaya
University of Barcelona and Institut d'Economia de Barcelona (IEB)

JANUARY 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Birds of a Feather Earn Together. Gender and Peer Effects at the Workplace*

Utilizing comprehensive administrative data from Brazil, we investigate the impact of peer effects on wages, considering both within-gender and cross-gender dynamics. Since the average productivity of both individuals and their peers is unobservable, we estimate these values using worker fixed effects while accounting for occupational and firm sorting. Our findings reveal that within-gender peer effects have approximately twice the influence of cross-gender peer effects on wages for both males and females. Furthermore, we observe a reduction in the disparity between these two types of peer effects in settings characterized by greater gender equality.

JEL Classification: J16, J24, J31, M12, M54

Keywords: peer effects, gender, matched employer-employee data, identity, wage determination

Corresponding author:
Anastasia Terskaya
Department of Economics
University of Barcelona
Gran Via de les Corts Catalanes
585, L'Eixample
08007 Barcelona
Spain

E-mail: a.terskaya@ub.edu

* We thank Pedro Albarrán, Marco Bertoni, Giorgio Brunello, Marco Caliendo, Dimitris Christelis, Lola Collado, Thomas Cornelissen, Sarah Pearlman, Núria Rodríguez-Planas, and Francesc Ortega as well as seminar participants at Universidad Carlos III de Madrid, LISER (Luxembourg), Queens College City University of New York, the University of Alicante, the University of Padova, IMT Lucca, the University of Valencia, the University of Navarra, Universidad Autónoma de Madrid, the 2021 Spring Meeting of Young Economists, the 2022 Jobs and Development Conference, the Royal Economic Society Annual Conference 2023, and the 2023 Networks and Development Workshop for their comments. Messina and Sanz-de-Galdeano acknowledge financial support from Project PID2021-124237NB-I00 (financed by MCIN/ AEI /10.13039/501100011033/ and by FEDER Una manera de hacer Europa) and from Generalitat Valenciana, Consellería de Innovación, Universidades, Ciencia y Sociedad Digital through project Prometeo CIPROM/2021/068. Terskaya acknowledges financial support from Projects PID2020-120589RA-I00 and PID2022-140330NB-I00 financed by MCIN/AEI.

1 Introduction

Does gender matter for peer effects in the workplace? Numerous multidisciplinary studies have underscored the key role of gender and gender social norms in shaping individual choices, social dynamics, and diverse economic outcomes. However, the influence of gender on peer effects at the workplace remains poorly understood. To fill in this gap, we exploit rich Brazilian matched employer-employee data to study whether the ability of same-gender peers has more influence on individual wages than that of cross-gender peers in a large local labor market.

The economic literature has long emphasized the role of peers in shaping productivity across diverse contexts, such as the workplace and school. In the workplace, individuals may conform to their peers' productivity due to interpersonal comparisons, effectively following a norm. Alternatively, these social interactions can foster mutual learning, commonly denoted as "knowledge spillovers", where workers enhance one another's productivity. Akerlof and Kranton (2000, 2002, 2005, 2008) proposed in a series of influential papers that identity can affect the interplay between norms and social interactions. In this context, work norms may well be gender-specific, leading to more relevant interactions among individuals who share a common gender identity. If gender identity indeed plays a substantial role in the workplace, individuals may be more strongly influenced by peers of the same gender than by those of the opposite gender.[1]

The objective of this paper is to assess the causal impact of the average permanent component of productivity of same-gender and opposite-gender peers on wages within an entire local labor market. We use comprehensive Brazilian matched employer-employee data encompassing over 7 million workers across more than 100,000

[1]The data used for the empirical analysis include binary information regarding workers' sex as indicated in their ID. Nevertheless, this information may not entirely align with their gender identity or their individual perception of gender. While the gender recorded on their ID is likely to match the gender identity of the majority of individuals, this may not hold true for everyone. Consequently, this situation introduces a potential source of measurement error in our gender variable, as discussed in Section 3. Throughout this paper, we use the term "gender" instead of "biological" or "legal sex" because gender identity is the primary determinant that influences the underlying mechanisms we are investigating.

formal firms in São Paulo city. Specifically, we consider a linear-in-means model (Manski, 1993) in which individual wages depend on the average productivity of both same-gender and opposite-gender coworkers from the same establishment and within the same narrowly defined occupational category.

We address several estimation challenges in our analysis. First, estimating peer effects requires a measure of peers' permanent component of productivity, which is unobservable. To tackle this issue, we leverage the panel structure of the data and use an iterative algorithm proposed by Arcidiacono et al. (2012), allowing spillovers to manifest through the fixed effects of coworkers from the same peer group. We adapt this algorithm to our setting to simultaneously estimate within-gender and cross-gender peer effect parameters and the permanent components of productivity of workers with worker fixed effects, which not only impact the wages of the focal worker but also the wages of their coworkers.

Second, estimating peer effects is complicated due to workers sorting into peer groups and contextual effects that might influence the wages of all workers within the same peer group, potentially introducing bias into the estimated peer effects. To address workers' sorting based on productivity, we control for worker fixed effects. To account for contextual effects, we also control for occupation-year, firm-year, and occupation-firm fixed effects.[2] Occupation-firm fixed effects take into account the possibility that highly productive occupation-firms may offer higher wages and attract highly productive workers. Occupation-year and firm-year effects control for wage fluctuations specific to particular occupations and firms that could be correlated with workers' productivity. Therefore, our primary assumption is that there are no time-varying peer-group-specific wage shocks correlated with the average workers' productivity after residualizing out our set of fixed effects. Via Monte Carlo simulations, we show that the potential bias introduced by time-varying peer group-level shocks, after adjusting for firm-year, occupation-year, and worker effects, is unlikely to be relevant enough to produce spurious results.

[2]Throughout the text, the term "firm" is often used to refer to the firm's plant or establishment.

One concern is that our results may be influenced by potential variations in job tasks between men and women, even when they share the same job titles. In our dataset, which is comprehensive, we address this concern by examining peer groups within 569 narrowly defined occupations, such as human resource managers, computer engineers, or IT administrators in our primary analysis. To further bolster the robustness of our results, we conduct a sensitivity check using an even more granular occupational classification, which distinguishes among 2,127 distinct occupations. This approach ensures that our conclusions remain uninfluenced by potential differences in job tasks performed by men and women within the same occupation.

Our main finding highlights the significantly greater influence of same-gender peers in the workplace compared to opposite-gender peers, although the latter also exert non-negligible effects for both men and women. In particular, the estimated wage elasticities with respect to the average productivity of same-gender and opposite-gender peers amount to 0.12-0.13 and 0.06-0.07, respectively, for both men and women. Furthermore, we explore various extensions of this primary empirical finding that indicate that the distinction between same-gender and opposite-gender peer effects becomes less pronounced in workplaces with greater gender equality.

First, we allow for heterogeneous effects across high- and low-skilled occupations. Many studies have indicated that traditional gender norms are more prevalent among less educated individuals, whereas highly educated individuals tend to hold more egalitarian views on gender roles (see Du et al. 2021 and the references therein). In line with this, our findings reveal that the difference between same-gender and opposite-gender peer effects is less pronounced in higher-skilled occupations than in lower-skilled ones.

Second, we explore heterogeneous effects among firms with varying gender wage gaps. We illustrate that the contrast between same-gender and opposite-gender effects is more subtle in firms with lower gender wage gaps, as opposed to firms with higher gender wage gaps. Finally, we show that the distinction between same-gender and opposite-gender peer effects is also less prominent in peer groups with a balanced composition of men and women compared to highly gender-segregated peer groups. However, it's worth noting that, even in contexts we categorize as gender-equal, there remains a positive disparity between same-gender and opposite-gender peer effects.

Our study contributes to several strands of the literature. First, we add to the body of literature examining peer effects on workers' output and productivity.[3] Our work is most closely related to Cornelissen et al. (2017), which studies peer effects on wages in a local labor market using worker-firm matched data for Germany.[4]. We adopt Cornelissen et al. (2017) estimation approach to assess peer effects in Brazil, and we further allow same-gender and opposite-gender peers to exert different influences on workers' productivity. Many prior studies of peer effects are based on laboratory experiments[5] or on real-world data from specific occupations such as cashiers (Mas and Moretti, 2009), farm workers (Bandiera et al., 2005, 2010; Brune et al., 2022), call center workers (Lindquist et al., 2015), or sportsmen (Guryan et al., 2009). None of these studies, however, distinguish between within-gender and cross-gender peer effects.

Second, our study contributes to a substantial body of literature that explores within-gender and cross-gender interactions in various social contexts. Prior research has established the role of same-gender role models, such as teachers or alumni speakers, in influencing educational and labor market outcomes, as well as lifetime well-being (Kofoed et al., 2019; Porter and Serra, 2020; Card et al., 2022; de Gendre et al., 2023; Patnaik et al., 2023). Another strand of this literature has investigated the effects of same-gender and opposite-gender peers on outcomes among teenagers, generally finding that same-gender influences tend to be more pronounced than opposite-gender peer effects on factors like high-school grades (Hsieh and Lin, 2017) or smoking behavior (Kooreman, 2007; Nakajima, 2007; Soetevent and Kooreman, 2007; Hsieh

[3]See Cornelissen (2016); Herbst and Mas (2015) and the references therein for a recent review of the literature.

[4]In a similar fashion, Hong and Lattanzio (2022) use matched employer-employee data for the region of Veneto to study peer effects on wages, and they additionally explore how these effects evolve over time and how the mobility of workers across firms affects coworkers wages in the origin and destination firms

[5]See for example Van Veldhuizen et al. (2018); Rosaz et al. (2016); Beugnot et al. (2019); Bellemare et al. (2010).

and Lin, 2017). Cools et al. (2022) explore the asymmetric effects of exposure during high school to female and male high flyers on various human capital and labor market outcomes. Our study expands upon this body of work by examining gender-specific peer effects in the context of the labor market.

Lastly, our study contributes to the extensive and growing body of literature that underscores the significant role of gender social norms in shaping economic outcomes, including those related to health and education (Rodríguez-Planas and Sanz-de Galdeano, 2019; Rodríguez-Planas et al., 2022; Guiso et al., 2008; Pope and Sydnor, 2010; Nollenberger et al., 2016; Anghel et al., 2020). These gender roles and attitudes exhibit variation across countries and have been shown to have a substantial influence on socioeconomic factors like fertility decisions, family formation, and female labor participation (Antecol, 2000, 2001; Fortin, 2005; Fernández and Fogli, 2006, 2009; Bertrand et al., 2015; Olivetti et al., 2020). Our research further enriches this literature by demonstrating how gender identity and associated social norms affect the influence of peers on wages, thus bridging the gap between this area of inquiry and the literature on peer effects.

The remainder of the paper is organized as follows. Section 2 explains our empirical strategy and discusses our identifying assumptions. Section 3 describes the data. Section 4 presents the results. Section 5 concludes.

2 Empirical Strategy

In this section, we lay out our empirical strategy, which is guided by a basic theoretical model that integrates aspects of identity economics into social interactions within the workplace. The proposed model illustrates the simplest form of peer effects where identity influences workplace dynamics. In this case, through peer pressure. The model can be extended to accommodate knowledge spillovers, and results are qualitatively similar. When knowledge spillovers or peer pressure are identity-dependent, same-gender peer effects on wages are more significant than opposite-gender effects.[6]

[6]The model including knowledge spillovers is available upon request.

2.1 Theoretical considerations

We outline the main components of the model, which is discussed in detail in Appendix A. We assume that worker i produces according to the following function:

$$f_{it} = \alpha_i + e_{it} + \epsilon_{it}, \tag{1}$$

where f_{it} is worker i's output at time t, α_i is individual i's ability (or his/her permanent component of productivity), e_{it} is individual i's effort, and ϵ_{it} is a random component of productivity independent of ability and effort. Because providing effort is costly but leads to higher wages, workers face a trade-off between earnings and effort on the job. In particular, there is a quadratic cost of providing effort:

$$C(e_{it}) = ce_{it}^2, \tag{2}$$

where $c > 0$ is a scale parameter.

In addition to the potential influence of peer pressure from coworkers on workers' utility, as documented in Kandel and Lazear (1992), Mas and Moretti (2009), and Cornelissen et al. (2017), our approach, consistent with the framework outlined by Akerlof and Kranton (2000, 2002, 2005, 2008), accommodates the notion that different individuals may experience varying degrees of peer pressure. Notably, the standards individuals, both men and women, must adhere to in order to mitigate disutility may be contingent on their gender. This gender-related asymmetry is explicitly accounted for in the subsequent peer pressure function:

$$P\Big(f_{it}, \mathbb{E}(f_{c_{it}}), \mathbb{E}(f_{\notin c_{it}})\Big) = (\eta + \eta_s)(\mathbb{E}(f_{c_{it}}) - f_{it})^2 + \eta(\mathbb{E}(f_{\notin c_{it}}) - f_{it})^2, \tag{3}$$

where $\mathbb{E}(f_{c_{it}})$ is the expected value of the production of workers who belong to the same social category as i and $\mathbb{E}(f_{\notin c_{it}})$ is the expected value of the production of workers who belong to a different social category than i.

The parameter $\eta \geq 0$ in equation (3) represents the pain of peer pressure that is independent of the social category of the worker. Instead, $\eta_s \geq 0$ represents potential asymmetries in the level of peer pressure. We assume that η_s depends on gender norms. If $\eta_s = 0$, gender norms are egalitarian, and all workers experience the same level of peer pressure from male and female peers, irrespective of their gender. However, if $\eta_s > 0$, individuals experience more peer pressure from peers belonging to the same social category than from different social category peers. The higher η_s is, the less egalitarian gender norms are.

Workers choose effort to maximize expected utility, as spelled out in Appendix A. Workers' output is observable and firms take wages as given, paying workers a piece rate that depends on workers' output. This simple setup leads to the following wage equation, which is the basis of our empirical analysis:

$$w_i = \kappa + \tilde{\alpha}_i + \tilde{\theta}_s \mathbb{E}(\tilde{\alpha}_{c_i}) + \tilde{\theta}_o \mathbb{E}(\tilde{\alpha}_{\notin c_i}) + v_i, \tag{4}$$

where κ is a positive constant that depends on the piece rate, $\tilde{\alpha}_i$ is a monetary representation of individual i's ability α_i, and $\tilde{\theta}_o$ and $\tilde{\theta}_s$ are two constants that depend on the parameters of the model.

Individual earnings are positively affected by workers' ability, but they also depend on the ability of peers of the same and opposite social categories. In Appendix A we show that

$$\tilde{\theta}_s - \tilde{\theta}_o = \frac{\eta_s}{c + 2\eta}, \tag{5}$$

which implies that the effect of same social category peers on wages is greater than the effect of opposite social category peers if and only if $\eta_s > 0$. If the social pressure function is symmetric ($\eta_s = 0$), same and opposite social categories have identical impacts on workers' wages, i.e., $\tilde{\theta}_o = \tilde{\theta}_s$. Note that $\tilde{\theta}_s - \tilde{\theta}_o$ is monotonically increasing in η_s. The less egalitarian gender norms are, the larger the gap in the response of workers' wages to peers of the same gender with respect to their response to peers of the opposite gender.

2.2 Empirical Specification

Equation (4) represents the conventional linear-in-means model of peer influence initially introduced by Manski (1993), extended to integrate insights from the economics of identity. In our context, there are two social categories, men and women. As we apply this equation to the data, we introduce additional flexibility across several dimensions. Specifically, our baseline empirical specification is as follows:

$$\begin{gathered} w_{itoj} = a_i + \alpha_1 \bar{a}^{f}_{\sim i,toj} Woman_i + \alpha_2 \bar{a}^{f}_{\sim i,toj} Man_i + \\ \beta_1 \bar{a}^{m}_{\sim i,toj} Woman_i + \beta_2 \bar{a}^{m}_{\sim i,toj} Man_i + \mu_{ot} + \rho_{jt} + \delta_{oj} + \varphi X'_{it} + v_{itoj}, \end{gathered} \tag{6}$$

where w_{itoj} is the (log) wage of individual i at time t in establishment j, and occupation o. The worker permanent component of productivity is denoted by a_i (the empirical counterpart of $\tilde{\alpha}_i$ from equation 4), while $\bar{a}^{f}_{\sim i,toj}$ and $\bar{a}^{m}_{\sim i,toj}$ represent the average permanent component of productivity of i's female and male peers, respectively (excluding i from the average in both cases). Note that, compared to equation (4), in the empirical specification we allow for differentiated effects of same-gender and opposite-gender peers across men and women. Thus, the coefficients $\alpha_1(\beta_1)$ and $\alpha_2(\beta_2)$ represent the impact on individual wages of the average permanent component of productivity of female(male) coworkers on women and men, respectively.

A worker's peer group is defined as all workers who work in the same establishment and occupation during the same year. Equation (6) includes occupation-year and establishment-year fixed effects (μ_{ot} and ρ_{jt}, respectively), as well as occupation-establishment fixed effect denoted by δ_{oj}. We will discuss below the reasons for including these fixed effects. Finally, X'_{it} is a vector of individual time-variant controls that include quadratic forms of age and firm tenure (the number of months individual i has been working in establishment j up to period t).

The estimation of peer effects in the workplace poses several challenges, some of which are generally applicable to all observational studies of peer effects (Manski, 1993). Additionally, we must address the fact that the workers' permanent component of productivity (a_i) is unobserved and needs to be estimated. We will discuss these issues sequentially.

2.3 Sorting and Omitted Variables

In our estimation, we follow Cornelissen et al. (2017) and condition on a large set of fixed effects to deal with sorting and omitted variables bias. We now outline the role played by each set of fixed effects for identification.

Workers are not randomly assigned and may self-select into peer groups. For instance, high-productivity workers may choose to join high-productivity peer groups. In such cases, the average productivity of peers ($\bar{a}^{m}_{\sim i,toj}$ and $\bar{a}^{f}_{\sim i,toj}$) and focal workers' wages (w_{itoj}) are likely to be positively correlated, even in the absence of any peer effect, because both variables are correlated with the productivity of worker i. This issue is addressed by controlling for each worker's permanent component of productivity (a_i), which we estimate using worker-fixed effects.

Additionally, even when we control for workers' permanent component of productivity, the average productivity of peers is likely correlated with other wage determinants, such as the quality of the occupation or establishment in which they work. Specifically, high-productivity peers may tend to be concentrated in highly productive occupations and/or establishments, which, in turn, are likely to offer higher wages. This may result in an overestimation of peer effects, as it would encompass not only peer effects but also the influence of omitted occupation-establishment characteristics. Therefore, it is necessary to control for occupation-establishment fixed effects (δ_{oj}) to account for time-invariant characteristics of occupations and establishments.

Furthermore, there may be occupation- or establishment-specific shocks that influence wages and attract different workers over time. For instance, establishments that adopt productivity-enhancing technologies may require a more highly skilled workforce and, as a result, offer higher wages. To account for time-varying establishment-specific and occupation-specific factors that may confound the analysis, we include occupation-year and establishment-year effects (μ_{ot} and ρ_{jt}, respectively).

To consistently estimate peer effects in equation (6), one must rely on the identifying assumption that there are no time-varying occupation-establishment-specific confounding factors that could influence wages and be correlated with the average productivity of coworkers, once controls have been partialled out. It's important to note that this identification also requires within-worker variation in $\bar{a}^{m}_{\sim i,toj}$ and $\bar{a}^{f}_{\sim i,toj}$. Such variation exists when: (i) worker i changes jobs, and (ii) new peers join or old peers leave their workplace. We exploit both sources of variation in the baseline empirical specification. Robustness checks discussed in Section 4.1 show that the empirical results are very similar when we eliminate individual worker mobility and rely exclusively on changes in the composition of his/her peers.

2.4 Estimation of Unobserved Productivity

The estimation of equation (6) is complicated because it involves the permanent component of worker i's productivity (a_i) and the average productivity of worker i's peers, $\bar{a}^{f}_{\sim i,toj}$ and $\bar{a}^{m}_{\sim i,toj}$, all of which are unobserved and need to be estimated. As a result, the model becomes a nonlinear least squares problem. Considering that we control for an extensive set of high-dimensional fixed effects (consisting of combinations of 7,583,748 workers, 111,110 establishment, 569 occupations, and 16 years), standard non-linear least squares techniques are not feasible. If $\bar{a}^{f}_{\sim i,toj}$ and $\bar{a}^{m}_{\sim i,toj}$ were observed instead, the model would be linear and could be estimated using conventional techniques.

To estimate equation (6) with the unobserved terms a_i, $\bar{a}^{f}_{\sim i,toj}$ and $\bar{a}^{m}_{\sim i,toj}$, we adapt the iterative algorithm proposed by Arcidiacono et al. (2012) to estimate spillover effects using panel data with both within- and cross-gender peer effects. We estimate peer effects by minimizing the following sum of squared residuals:

$$\min_{a,\alpha,\beta,\rho,\delta,\mu,\varphi} \sum_{i=1}^{N} \sum_{t=1}^{T} \Big(w_{itoj} - a_i - \alpha_1 \bar{a}^{f}_{\sim i,toj} Woman_i - \alpha_2 \bar{a}^{f}_{\sim i,toj} Man_i - \\ \beta_1 \bar{a}^{m}_{\sim i,toj} Woman_i - \beta_2 \bar{a}^{m}_{\sim i,toj} Man_i - \mu_{ot} - \rho_{jt} - \delta_{oj} - \varphi X'_{it} - v_{itoj} \Big)^2 \tag{7}$$

The algorithm proceeds as follows:

1. Set an initial guess for the vector of fixed effects (a) a^0.

2. Conditional on a^0, compute $\bar{a}^f_{\sim i,toj}$ and $\bar{a}^m_{\sim i,toj}$ and estimate $\alpha_1, \alpha_2, \beta_1, \beta_2$, and the rest of the parameters (μ_{ot}, ρ_{jt},δ_{oj},φ) by OLS.

3. Update a^1 according to equations (B.3) and (B.4), the first order conditions for (7) derived in Appendix B.

4. Iterate steps 2 and 3 until convergence of $\alpha_1, \alpha_2, \beta_1$, and β_2 is achieved.

Convergence is achieved if the sum of squared residuals diminishes with every iteration, which requires the right-hand side of equations (B.3) and (B.4) to be a contraction mapping. The following theorem provides sufficient conditions for convergence.

Theorem 1 *Denote N_w as the number of women and N_m as the number of men. Denote $g^w(a) : \mathbb{R}^{N_w} \to \mathbb{R}^{N_w}$ and $g^m(a) : \mathbb{R}^{N_m} \to \mathbb{R}^{N_m}$, where the i^{th} element of $g^f(a)$ is given by the right-hand side of* (B.3) $\forall i \in N_w$*, and the i^{th} element of $g^m(a)$ is given by the right-hand side of* (B.4) $\forall i \in N_m$*. $g^f(a)$ and $g^m(a)$ are contraction mappings if $\alpha_k < 0.2$ and $\beta_k < 0.2$ for $k = 1, 2$.*

The proof of this Theorem is provided in Appendix B. Intuitively, Theorem 1 suggests that convergence may not be achieved when peer effects are exceedingly large.

In the algorithm proposed by Arcidiacono et al. (2012), the estimated peer effects are $\sqrt{N}$ consistent and asymptotically normal estimators if the residuals (ν_{itoj} in equation 6) between any two observations are uncorrelated. This assumption eliminates both serial correlation and the presence of any wage shocks shared among the peer group. This may be considered an excessively restrictive assumption in our context. We run a series of Monte Carlo simulations, discussed in Section 4.1 and described in detail in Appendix C, to assess the sensitivity of the estimator to violations of this assumption. We conclude that the bias in the peer effect estimates due to serial correlation of a plausible magnitude is modest. Moreover, time-varying peer group-level shocks may introduce an upward bias, but this bias is not substantial enough to spuriously generate the magnitude of the peer effects that we estimate.

3 Data

We utilize data from Brazil's Relação Anual de Informações Sociais (RAIS), an administrative dataset that matches employers and employees, collected annually by the Brazilian Ministry of Labor. Data collection began in 1975, but we are using data from 2003 to 2018 because the occupational classification changed in 2003 and remained consistent in subsequent years. RAIS data is derived from a mandatory survey filled out by all formally registered firms in Brazil, providing information on earnings and demographic characteristics of workers as reported by their employers. According to estimates from the Ministry of Labor, it covers approximately 98% to 99% of all officially registered firms.

RAIS includes demographic information about workers, such as their biological sex, age, and education, as well as details about their jobs, including contract type, occupation, average monthly earnings for the year, tenure within the firm, and typical weekly working hours.[7] Additionally, RAIS contains certain characteristics of the firms' establishments, such as their sector, region, and municipality. Notably, RAIS also provides unique and anonymized identifiers for firms, establishments, and individual workers. These identifiers, along with the dates of employment entry and separation, enable us to track workers and employers over time and distinguish between job stayers and switchers.

Our dependent variable in the regressions is the log of monthly wages. If workers change employers at any point in time during the calendar year, their monthly wage refers to the average monthly wage during the employment spell. Wages in RAIS are expressed as multiples of the minimum wage in December of that year. To convert the wage variable into real values, we adjust it by multiplying it by the national minimum wage in December of that year, which has been deflated using the consumer price index.

[7]The fact that RAIS contains information about workers' biological sex rather than their gender could potentially introduce measurement errors when forming same-gender and opposite-gender peer groups. These errors may, in turn, attenuate the distinctions between same-gender and opposite-gender peer effects, leading to the interpretation of these differences as a lower bound of the true discrepancy.

3.1 Sample Selection and Peer Group Definition

We use data from 2003 to 2018 for the city of São Paulo. By focusing on one large local labor market rather than a random sample of all formal workers in Brazil, we aim to capture the majority of worker mobility, which is essential for our identification strategy. Our sample consists of full-time workers, defined as those working at least 30 hours per week, in the private sector, aged between 15 and 65, with valid information regarding gender, wages, tenure, occupation, and firms' establishment identifiers.

We define a worker's peer group as all individuals employed in the same establishment and the same four-digit occupation during the same year. Specifically, occupations are categorized according to the Brazilian Classification of Occupations, known as CBO 2002.[8] CBO 2002 comprises 620 four-digit occupations, such as human resource managers, computer engineers, IT administrators, nutritionists, lawyers, telephone operators, and home sellers. Occupational definitions at the four-digit level are considered sufficiently detailed to ensure that workers can potentially interact, observe, and assess their peers' performance. However, we also conduct a robustness check in Section 4.1, where we use six-digit occupation codes. For instance, within the four-digit occupation category of "lawyer", a six-digit code distinguishes between various types of lawyers, including corporate lawyers such as civil lawyers, public lawyers, criminal lawyers, specialized lawyers, labor lawyers, and legal consultants. Consequently, if men and women sort into different narrow occupations within their four-digit occupation, this variation is likely captured by the six-digit occupation code. In total, we can distinguish 2,127 six-digit occupations.

To investigate whether there are systematic gender differences in employment within the peer groups, which are initially defined using four-digit occupations, we examine gender differences in occupational characteristics within these peer groups. Specifically, we calculate the average wages and the share of women at the six-digit occupation level. Subsequently, we regress each of these variables on a female dummy variable and the vector of peer group (defined at the four-digit level) fixed effects.

[8] The classification is available at `http://www.mtecbo.gov.br./cbosite/pages/home.jsf`. Prior to 2003, occupations were classified according to CBO 1994.

Our findings indicate that women are employed in occupations with wages that are 0.09 percent lower, and that there is a 0.28 percentage point higher representation of women compared to men within the same four-digit level peer group. These differences are minimal, providing assurance that there are no significant systematic variations in the tasks performed by men and women classified as peers in our benchmark four-digit level definition.

Job spells can begin and end in any month throughout the year. Therefore, to confirm that the members of the peer group have indeed been in the same workplace and occupation at some point during the year, we only include workers who were employed in that specific occupation and establishment in November. Additionally, when individuals hold multiple jobs, we retain the observation corresponding to the highest-paying job. Applying these filters results in a sample comprising 57,726,566 worker-year observations.

To estimate equation (6), it is necessary that each employee has at least one male peer and one female peer. Consequently, we exclude all peer groups with fewer than two males and two females. This additional filter reduces our sample size to 27,510,560 observations. Finally, we narrow our analysis to the largest connected mobility group because fixed effects are only identifiable within establishments directly or indirectly connected by worker mobility over the entire sample period. This results in a final sample size of 27,464,523 worker-year observations.

3.2 Descriptive Statistics

Table 1 provides an overview of the panel structure of our estimation sample. The sample comprises 27,464,523 worker-year observations from 7,583,748 unique formal employees, 111,110 establishments, 569 four-digit occupations, and 913,953 peer groups (occupation-establishment-year combinations). The presented statistics indicate a notable degree of worker mobility across establishments and occupations, which is crucial for identifying worker, establishment-occupation, establishment-year, and occupation-year fixed effects. Specifically, over our sample period (2003-2018), workers are observed for an average of 3.61 years, working for an average of 1.6 different establishments and in 1.5 different occupations.

Table 2 presents additional descriptive statistics split by gender. Women represent 49.3% of our sample and earn, on average, 36.2% less than men. Additionally, women in our sample have a higher average level of education compared to male workers, as the proportion of women with only primary education is lower than that of men. This suggests that there may be selection into employment based on both education and gender. We account for this selection by including worker fixed effects in equation (6).

Implementing our identification strategy requires the presence of various types of data variation. First, we require a sufficient number of observations within each peer group. Panel B of Table 2 shows that, on average, peer groups consist of 30 workers, with a median of 10 workers. Second, we rely on within-peer-group wage variation. The data in Panel C of Table 2 clearly indicates that there is substantial wage variability among employees within the same peer group. Specifically, the average within-peer-group standard deviation of log real wage residuals amounts to 0.382, which is equivalent to 45% of the overall standard deviations in log real wages (0.849).

As discussed in Section 2.3, identifying the wage effects of peers' permanent component of productivity in equation 6 requires the presence of within-worker variation in peers' composition. Such variation arises when workers change jobs or when their peer group composition changes due to the entry or exit of some coworkers. In this context, Table 3 presents an overview of the extent of within-worker variation in peers' composition for both men and women. The standard deviation of the annual change in the average peers' fixed effect is 0.18 for men and 0.16 for women, which is approximately 30% of the standard deviation of the average fixed effect of male peers ($\bar{a}^{m}_{\sim itoj}$) and female peers ($\bar{a}^{w}_{\sim itoj}$). As expected, the standard deviation of the annual change in the average peers' fixed effect is about three times larger among job movers than among job stayers, reflecting the greater turnover and composition changes experienced by job movers. However, within-worker variation of the peers' average fixed effect among job stayers is not insignificant, accounting for approximately 20% of the standard deviation of the peers' average fixed effect. In Section 4.1, we conduct a separate analysis of peer effects for movers and stayers.

Table 3 also shows that the correlation between individual workers' fixed effects and the average fixed effect of their peers is approximately 0.8 for both male and female peers. This suggests that high-quality employees tend to sort into high-quality peer groups, underscoring the importance of accounting for worker fixed effects in equation 6. Furthermore, we observe positive correlations between the average fixed effect of peers and establishment-occupation effects (δ_{oj}), establishment-year effects (ρ_{jt}), and occupation-year effects (μ_{ot}). This suggests that high-quality peer groups tend to be concentrated in highly productive and high-wage occupation-establishment combinations, stressing the importance of including occupation-establishment effects, establishment-time effects, and occupation-time effects in equation (6).

4 Results

We now turn to our main question of interest: do same-gender and opposite-gender peers have the same impact on individual wages? To answer this question, we estimate equation (6) controlling for the full set of controls and fixed effects described above.

Figure 1 illustrates the estimated peer effects for both men and women. Our findings reveal that the influence of the permanent productivity component of same-gender peers on the wages of focal workers, regardless of their gender, is roughly twice as large as that of opposite-gender peers. Specifically, a 10% increase in the average fixed effect of female peers results in a 1.34% increase in female wages, which is approximately twice the wage increase observed when the average fixed effect of male peers increases by 10%. Similarly, a 10% increase in the average fixed effect of male peers leads to a 1.21% increase in male wages, whereas a 10% increase in the quality of female peers has a 0.74% impact on male wages. The point estimates and their corresponding standard errors are reported in Table 4.

The magnitude of our estimated peer effects falls within the range of peer effects on focal worker output reported in previous laboratory experiments and field studies conducted in specific settings. On average, these studies found peer effects of approximately 0.12 (SE=0.03), as documented in the meta-analysis by Herbst and Mas (2015). Interestingly, our estimated peer effects are considerably larger than the 0.1% wage increase induced by a 10% increase in peers' quality estimated by Cornelissen et al. (2017) using matched employer-employee data for an entire local labor market in Germany.[9]

A potential explanation for the observed difference between same-gender and opposite-gender peer effects is the relevance of gender identity norms in the workplace. Specifically, the literature on identity economics suggests that individuals may be inclined to conform to the behaviors exhibited by their peers within their social category, such as gender (Akerlof and Kranton, 2000, 2002, 2005, 2008). This conformity arises from the belief that adhering to gender-specific behavioral norms reinforces and validates one's self-image or identity. As a result, men and women may experience more pressure to match the productivity of their male and female coworkers, respectively, which aligns with the productivity spillovers we have uncovered. Furthermore, if interactions within the same gender category are more prevalent than those across gender categories, individuals may learn more from peers of the same gender than from peers of the opposite gender.

4.1 Sensitivity Analyses

To assess the robustness of the findings in Figure 1, we conduct a set of sensitivity analyses.

[9]The factors contributing to cross-country variations in workplace peer effects are outside the scope of this paper. Nevertheless, it is pertinent to highlight two distinctions between Germany and Brazil that may be relevant to this comparison. First, Germany features a considerably more unionized labor market than Brazil. Thus, we should expect more individual wage bargaining in Brazil, potentially leading to more pronounced peer effects. Second, Brazil lags behind developed countries like Germany in terms of the availability of formal training programs. In this context, the impact of peers on learning-by-doing and the assessment of inter-group productivity comparisons may be more significant.

Movers *versus* stayers. We can identify peer effects from two distinct sources of variation: i) changes in peers' composition caused by focal workers leaving their jobs (job movers); and ii) changes in peers' composition due to the entry or exit of peers for job stayers. As depicted in Table 3, there exists variation along both dimensions, but, as anticipated, the variation among job movers is more pronounced than among stayers.

To mitigate concerns related to the endogeneity of mobility decisions in the presence of match-specific effects, we replicate our main analysis, as presented in Figure 1, by separately examining peer effects among job movers and job stayers. The latter group accounts for match-specific effects. The results of this analysis are detailed in Panels A and B of Table 5. Reassuringly, we find that peers' average fixed effects influence the wages of focal workers for both movers and stayers, with the impact of same-gender peers being stronger than that of opposite-gender peers. Notably, for both movers and stayers, our estimates closely resemble the baseline estimates outlined in Table 4, indicating that these are not biased by match-specific effects. [10]

Narrowing the peer group definition In any peer effects study, it is imperative to thoroughly assess the appropriateness of the chosen peer group definition. In our context, the critical question is whether workers in a specific occupation, as defined at the four-digit level within an establishment, genuinely interact, have opportunities to learn from their peers, observe and evaluate their productivity, and potentially react to it. If the peer group definition is excessively narrow, leading to the exclusion of relevant peers, our estimates are likely to be attenuated. If, instead, the peer group definition is excessively broad, it could be the case that the distinctions between same-gender and opposite-gender peer effects we have uncovered thus far merely mask the fact that men and women often perform distinct tasks within the same four-digit occupation at the same establishment.

In Panel C of Table 5, we replicate our benchmark analysis utilizing a more re-

[10]It's also worth noting that Cornelissen et al. (2017) report similar results (see their Table 7), and in accordance with this evidence, Card et al. (2013) and Card et al. (2015) find that idiosyncratic job-match effects are not a significant driver of job mobility.

strictive peer group definition: coworkers in the same establishment within the same occupation defined at the six-digit level. The six-digit level occupational classification reflects a detailed occupational profile that precisely delineates the range of tasks performed. After limiting the analysis to peer groups defined as six-digit level occupations, each containing a minimum of 2 male and 2 female workers, the estimation sample includes 2,127 six-digit occupations, and is reduced to 25,578,603 observations.

The results reveal that the estimated peer effects obtained when defining peer group at the six-digit occupational level are similar to our four-digit level benchmark. The estimated effects of same-gender peers are about 1.4 and 1.8 times larger than the effects of opposite-gender peers for men and women, respectively.

Gender-balanced occupations In occupational settings where there is significant gender segregation, male employees are more likely to interact predominantly with other men, and female employees with other women. If such segregation remains constant over time, these gender imbalances will be accounted for by the occupation-establishment fixed effects introduced in equation (6). However, a concern arises if our findings are influenced by occupation-establishments that are becoming increasingly segregated by gender over time.

To evaluate the impact of gender segregation in establishment-occupations, we analyzed peer effects in gender-balanced occupation-establishments, identified as those with a male and female representation ranging from 40% to 60%. These results are presented in Panel D of Table 5. We found that the influence of same-gender peers on individual wages is more pronounced than that of opposite-gender peers for both genders. This highlights that our main conclusions are not merely a result of the varying distribution over time of men and women across different occupation-establishments.

Gender-specific fixed effects and group-level controls. While we control for a set of fixed effects that account for firm and occupation wage-relevant characteristics, it could be the case that firms and occupations affect men and women's wages differently. To account for this, in Panel E of Table 5 we estimate the model with gender-specific fixed effects. That is, we estimate equation (6) with occupation-year-gender, establishment-year-gender, and occupation-establishment-gender fixed effects. The estimated effects are similar to the main results reported in Figure 1 and indicate that same-gender peer effects are more relevant than opposite-gender peer effects for both men and women.

Next, we test the sensitivity of our results by including peer-group-specific controls. In Panel F of Table 5, we estimate equation (6) while accounting for factors such as the share of women, average age, average tenure within the peer group, and the size of the peer group. The estimated effects closely resemble the baseline results.

Excluding small or large peer groups. Next, we test whether our results change if we exclude large and small peer groups. In Panel G of Table 5 we focus on peer groups with 10-50 workers. The results are similar to the results obtained with our benchmark specification.

Peer effects estimated for all of Brazil. In our main analysis, we focus on the city of São Paulo instead of using data for the entire Brazilian labor market to ease the computational burden. São Paulo serves as the industrial hub in Brazil, and the results may not be fully generalizable to the entire country. Therefore, we also estimate our main specification for all of Brazil.

For this estimation, we utilize a sample of 200,253,593 worker-year observations involving 49,987,482 workers in 1,047,376 establishments after applying similar filters as in our baseline sample. The estimated effects are provided in Panel H of Table 5. The results closely resemble the estimates for São Paulo and reveal that same-gender effects are stronger than opposite-gender effects for both men and women. If anything, peer effects are larger, and the differences between same-gender and opposite-gender peer effects are exacerbated.

Placebo Tests We also conducted two placebo tests to address the concern that our results might be due to chance. First, we estimate the effects of the average fixed effect of workers in other randomly selected firms within the same four-digit occupation. It's important to note that the effects of peers from other establishments need not be zero, as knowledge spillovers can occur between employees from different establishments, particularly if they are geographically and economically close (Moretti, 2004). However, it is expected that peer effects will be stronger among peers within the same establishments than among employees in the same occupation working in different, randomly chosen establishments.

Consistent with this expectation, the average fixed effect of workers in the same four-digit occupation but working in a different establishment has no discernible influence on the wages of focal workers. Panel I of Table 5 displays the results, demonstrating that the estimated placebo peer effects never exceed a magnitude of 0.002.

Second, we estimate the effects of the average fixed effect of workers in randomly chosen four-digit occupations within the same establishment. Consistent with our main results being genuine, we find no evidence of peer effects across occupations within the same establishment (Panel J, Table 5).

Monte Carlo Simulations To achieve a $\sqrt{N}$ consistent and asymptotically normal estimator of peer effects for a fixed T, it is necessary to assume that the residuals between any two observations are uncorrelated, as outlined in Theorem 1 in Arcidiacono et al. (2012). The presence of random shocks that affect workers in the same peer group would violate this assumption. Similarly, serial correlation in the individual error term would also violate this assumption.

To assess the extent of bias associated with (i) peer group-specific shocks and (ii) serial correlation in the individual error term, we conduct Monte Carlo simulations. Additional details about the simulation can be found in Appendix C

The results of these simulations are provided in Table C.1 in Appendix C. Biases are calculated as the difference between the estimated effects and the coefficients assumed in the data-generating process. The results indicate that a peer group-specific shock is associated with an upward bias of 0.003-0.008 when the variance of the peer-group shock, as a share of the total error variance, is assumed to be approximately

3%. This bias remains of a similar magnitude regardless of the assumed value of the true coefficients.

When assuming that the peer-group-specific shock, as a share of the total error variance, is 6%, the bias ranges from 0.009 to 0.016. However, this bias is small in comparison to the estimated peer effects we obtain. This suggests that it is unlikely that our estimates are significantly influenced by peer group-specific shocks that are not accounted for by firm and occupation time-variant fixed effects.

Finally, the results presented in Table C.1 in Appendix C also indicate that serial correlation of a plausible magnitude is unlikely to have a significant biasing effect on our estimates.

4.2 Are Differences Between Same-Gender and Opposite-Gender Peer Effects Exacerbated in Less Gender Equal Contexts?

We will now investigate whether the disparities between peer effects among individuals of the same gender and those of opposite genders are less prominent in contexts characterized by greater gender equality, as posited in the model presented in Section 2.2. This is in line with our hypothesis that a reduction in gender-specific behavioral expectations would alleviate the sense of identity loss experienced by women and men who engage in behaviors that are more commonly associated with the opposite gender. To test this hypothesis, we will estimate our model across various subgroups. More specifically, we will partition the sample based on occupational skill levels, the establishment-level gender wage gap, and the extent of gender segregation in the workplace.

Figure 2 presents the estimated effects in various subsamples. Overall, the difference between the same-gender and the opposite-gender peer effects is positive across all contexts, but it tends to be more pronounced in less egalitarian job settings. The estimated effects, along with their corresponding standard errors are presented in Table 6.

Peer Effects by occupational skill level. Panels A and B of Figure 2 depict the peer effects categorized by occupational skill level, which we define as the bottom/top 50% of occupations with the lowest/highest share of employees with post-secondary education. In general, we observe that peer effects are more pronounced in low-skilled occupations compared to high-skilled ones.[11] This finding aligns with the results presented in Cornelissen et al. (2017), highlighting peer pressure as the primary driver of workplace peer effects, as opposed to knowledge spillovers.[12]

Numerous studies on gender identity have consistently shown that traditional gender norms are more prevalent among individuals with lower levels of education, whereas those with higher education tend to endorse more egalitarian views regarding the roles of men and women in society (see Du et al. 2021; Rivera-Garrido 2022 and references therein). Brazil follows this pattern, as highly educated respondents in the World Value Survey from Brazil are less likely to adhere to traditional gender norms than their less educated counterparts.[13] Consequently, one would anticipate that the disparities in peer effects between individuals of the same gender and those of the opposite gender would be more pronounced in low-skilled occupations than in high-skilled ones.

The results presented in Panels A and B of Figure 2 confirm this observation for men. In high-skilled occupations, the same-gender peer effect for men is 1.5 times larger than the opposite-gender peer effect, whereas in low-skilled occupations, this ratio increases to 2.9. In contrast, for women, the difference between same-gender and opposite-gender peer effects remains similar in both high-skilled and low-skilled occupations.

[11]We also attempted alternative partitions, such as quartiles. However, peer effects at the lowest quartile of the skill distribution appear to be larger than 0.2, and the conditions for convergence of the estimator were not met.

[12]For a detailed discussion, refer to Section IV.B of Cornelissen et al. (2017).

[13]Analysis of World Value Survey data for Brazil from 2010-2014 reveals negative correlations between completing college education and agreeing with statements reflecting traditional gender norms, such as (i) *"It is a problem if women have more income than their husbands"* and (ii) *"When jobs are scarce, men should have more right to a job than women"*.

Peer effects by establishment-level gender wage gap Previous studies suggest that gender wage gaps tend to be more pronounced in contexts characterized by less egalitarian gender identity norms (see Antecol 2001; Fortin 2005).[14] In line with this concept, we use establishment-level gender wage gaps (adjusted for variations in the composition of education and age across establishments) as a proxy for establishment-level gender equality.[15] More precisely, we estimate same-gender and cross-gender peer effects for both men and women within the lower and upper quartiles of the firm-level gender wage gap distribution.

Panels C and D in Figure 2 and Table 6 display the results of our analysis. They reveal that women working at establishments with a low gender wage gap are influenced similarly by both male and female coworkers. In contrast, women in firms with a high wage gender gap are 2.5 times more influenced by their female coworkers than their male coworkers. For men, the difference in peer effects between same-gender and opposite-gender interactions is similar in both high- and low-wage-gender-gap firms.

Peer effects by establishment-level differences in female leadership The global trend of women's underrepresentation in managerial roles is evident across various countries, including Brazil, as highlighted in (OECD/ILO, 2021). We suggest that organizations with a more gender-inclusive corporate culture tend to appoint more women to leadership positions. This, in turn, may reduce differences in peer effects observed within the same gender compared to those across different genders.

Panels E and F of Figure 2 and Table 6 provide additional insights. These analyses focus on establishments with varying proportions of women in the top 25% of their wage distributions. The results support our previous findings, showing a smaller

[14]Common sources of gender pay gaps, including differences in working hours and participation (Olivetti and Petrongolo, 2008), fertility (Kleven et al., 2019), and labor market institutions (Blau and Kahn, 2003) are unlikely to be relevant in our context, which compares gender pay gaps across firms using full-time workers within the local labor market of Sao Paulo.

[15]To calculate the establishment-level gender wage gap, we follow these steps: Separately for men and women, we perform regressions of wages (in logarithm) on educational level, age, age-squared, and year fixed effects. Subsequently, we calculate the establishment-level means of the residuals of these regressions for both men and women. The gender gap is calculated as the difference between the mean residuals of men and women.

gap between same-gender and opposite-gender peer effects in establishments within the top quartile for female representation among top earners. In contrast, a significantly larger gap is observed in establishments in the bottom quartile. Consistent with earlier observations, the disparity in peer effects across more egalitarian and less egalitarian contexts affects women more than men. In environments with fewer women in leadership, the productivity of female peers has almost triple the impact on women's wages compared to male peers' productivity. Conversely, in establishments with a high proportion of women in top positions, the differential impact of male versus female peer effects on women's wages is minimal.

5 Conclusion

We present new evidence that underscores the presence of robust peer effects in the workplace, wherein same-gender effects are more prominent than opposite-gender effects. Leveraging longitudinal matched employer-employee data from Brazil, our investigation unveils that the influence of same-gender peers' permanent productivity component on the wages of focal workers is approximately twice as large as the influence of opposite-gender peers, a phenomenon observed among both female and male workers. Our findings remain highly robust across a comprehensive battery of checks, and the observed asymmetries cannot be attributed to disparities in job responsibilities among men and women within the same occupations and establishments, nor to gender-based job sorting

This evidence supports the idea that gender identity norms exert an influence on workplace interactions. In line with this concept, we also observe that in contexts characterized by greater gender equalitysuch as high-skilled occupations or establishments with narrower gender wage gapsthe disparities between same-gender and opposite-gender peer effects are less prominent.

Our findings contribute to the understanding of the determinants of gender inequalities in the labor market and their persistence. The observed disparities between same-gender and opposite-gender peer effects suggest that gender imbalances in the workplace may be further exacerbated due to the social multiplier effects on wages. This observation aligns with arguments presented by Arduini et al. (2020a) and Arduini et al. (2020b), who underscore the importance of accounting for heterogeneity in peer effects when devising and assessing policy interventions. In cases where an intervention has a more substantial positive impact on one group compared to another, inequality in outcomes between the two groups could potentially be intensified in contexts where within-group peer effects are robust, and between-group peer effects are less influential.

Consider, for instance, the implications of our findings for on-the-job training participation. It is well-established that changes in worker productivity resulting from on-the-job training can influence the productivity of non-trained coworkers (see Lindquist et al. 2015; De Grip and Sauermann 2012). Our study suggests that any imbalances in training uptake between men and women would have multiplier effects through peer effects, with implications for the gender wage gap.[16] A company aiming to reduce the gender wage gap might, therefore, incorporate specific incentives for the training of women into its gender strategy.

References

Akerlof, G. A. and R. E. Kranton (2000): "Economics and identity," *The Quarterly Journal of Economics*, 115, 715–753.

——— (2002): "Identity and schooling: Some lessons for the economics of education," *Journal of Economic Literature*, 40, 1167–1201.

——— (2005): "Identity and the Economics of Organizations," *Journal of Economic perspectives*, 19, 9–32.

——— (2008): "Identity, supervision, and work groups," *American Economic Review*, 98, 212–17.

[16] Although international evidence is scarce, these imbalances appear to exist in practice. For instance, Caliendo et al. (2022) shows that women were significantly less likely than men to participate in firm-specific training in Germany during the 2000s.

ANGHEL, B., N. RODRÍGUEZ-PLANAS, AND A. SANZ-DE GALDEANO (2020): "Is the math gender gap associated with gender equality? Only in low-income countries," *Economics of Education Review*, 79, 102064.

ANTECOL, H. (2000): "An examination of cross-country differences in the gender gap in labor force participation rates," *Labour Economics*, 7, 409–426.

——— (2001): "Why is there interethnic variation in the gender wage gap?: The role of cultural factors," *Journal of Human Resources*, 119–143.

ARCIDIACONO, P., G. FOSTER, N. GOODPASTER, AND J. KINSLER (2012): "Estimating spillovers using panel data, with an application to the classroom," *Quantitative Economics*, 3, 421–470.

ARDUINI, T., E. PATACCHINI, AND E. RAINONE (2020a): "Identification and estimation of network models with heterogeneous interactions," in *The Econometrics of Networks*, Emerald Publishing Limited.

——— (2020b): "Treatment effects with heterogeneous externalities," *Journal of Business & Economic Statistics*, 38, 826–838.

BANDIERA, O., I. BARANKAY, AND I. RASUL (2005): "Social preferences and the response to incentives: Evidence from personnel data," *The Quarterly Journal of Economics*, 120, 917–962.

——— (2010): "Social incentives in the workplace," *The Review of Economic Studies*, 77, 417–458.

BELLEMARE, C., P. LEPAGE, AND B. SHEARER (2010): "Peer pressure, incentives, and gender: An experimental analysis of motivation in the workplace," *Labour Economics*, 17, 276–283.

BERTRAND, M., E. KAMENICA, AND J. PAN (2015): "Gender identity and relative income within households," *The Quarterly Journal of Economics*, 130, 571–614.

BEUGNOT, J., B. FORTIN, G. LACROIX, AND M. C. VILLEVAL (2019): "Gender and peer effects on performance in social networks," *European Economic Review*, 113, 207–224.

BLAU, F. D. AND L. M. KAHN (2003): "Understanding International Differences in the Gender Pay Gap," *Journal of Labor Economics*, 21, 106–144.

BRUNE, L., E. CHYN, AND J. KERWIN (2022): "Peers and motivation at work: evidence from a firm experiment in Malawi," *Journal of Human Resources*, 57, 1147–1177.

CALIENDO, M., D. A. COBB-CLARK, C. OBST, H. SEITZ, AND A. UHLENDORFF (2022): "Locus of control and investment in training," *Journal of Human Resources*, 57, 1311–1349.

CARD, D., A. R. CARDOSO, AND P. KLINE (2015): "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women," *The Quarterly Journal of Economics*, 131, 633–686.

Card, D., C. Domnisoru, S. G. Sanders, L. Taylor, and V. Udalova (2022): "The Impact of Female Teachers on Female Students' Lifetime Well-Being," Working Paper 30430, National Bureau of Economic Research.

Card, D., J. Heining, and P. Kline (2013): "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly Journal of Economics*, 128, 967–1015.

Cools, A., R. Fernández, and E. Patacchini (2022): "The asymmetric gender effects of high flyers," *Labour Economics*, 79, 102287.

Cornelissen, T. (2016): "Do social interactions in the workplace lead to productivity spillover among co-workers?" *IZA World of Labor*, 107.

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

de Gendre, A., J. Feld, N. Salamanca, and U. Zölitz (2023): "Same-Sex Role Model Effects in Education," Working Paper 438, University of Zurich.

De Grip, A. and J. Sauermann (2012): "The effects of training on own and co-worker productivity: Evidence from a field experiment," *The Economic Journal*, 122, 376–399.

Du, H., Y. Xiao, and L. Zhao (2021): "Education and gender role attitudes," *Journal of Population Economics*, 34, 475–513.

Fernández, R. and A. Fogli (2006): "Fertility: The role of culture and family experience," *Journal of the European Economic Association*, 4, 552–561.

——— (2009): "Culture: An empirical investigation of beliefs, work, and fertility," *American Economic Journal: Macroeconomics*, 1, 146–77.

Fortin, N. M. (2005): "Gender role attitudes and the labour-market outcomes of women across OECD countries," *Oxford Review of Economic Policy*, 21, 416–438.

Guiso, L., F. Monte, P. Sapienza, and L. Zingales (2008): "Culture, gender, and math," *Science*, 320, 1164–1165.

Guryan, J., K. Kroft, and M. J. Notowidigdo (2009): "Peer effects in the workplace: Evidence from random groupings in professional golf tournaments," *American Economic Journal: Applied Economics*, 1, 34–68.

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

Hong, L. and S. Lattanzio (2022): "The Peer Effect on Future Wages in the Workplace," *Available at SSRN 4052587*.

Hsieh, C.-S. and X. Lin (2017): "Gender and racial peer effects with endogenous network formation," *Regional Science and Urban Economics*, 67, 135–147.

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

KLEVEN, H., C. LANDAIS, J. POSCH, A. STEINHAUER, AND J. ZWEIMULLER (2019): "Child Penalties across Countries: Evidence and Explanations," *AEA Papers and Proceedings*, 109, 122–126.

KOFOED, M. S. ET AL. (2019): "The effect of same-gender or same-race role models on occupation choice evidence from randomly assigned mentors at west point," *Journal of Human Resources*, 54, 430–467.

KOOREMAN, P. (2007): "Time, money, peers, and parents; some data and theories on teenage behavior," *Journal of Population Economics*, 20, 9–33.

LINDQUIST, M. J., J. SAUERMANN, AND Y. ZENOU (2015): "Network effects on worker productivity," *CEPR Discussion Paper No. DP10928*.

MANSKI, C. F. (1993): "Identification of endogenous social effects: The reflection problem," *The Review of Economic Studies*, 60, 531–542.

MAS, A. AND E. MORETTI (2009): "Peers at work," *American Economic Review*, 99, 112–45.

MORETTI, E. (2004): "Workers' education, spillovers, and productivity: evidence from plant-level production functions," *American Economic Review*, 94, 656–690.

NAKAJIMA, R. (2007): "Measuring peer effects on youth smoking behaviour," *The Review of Economic Studies*, 74, 897–935.

NOLLENBERGER, N., N. RODRÍGUEZ-PLANAS, AND A. SEVILLA (2016): "The math gender gap: The role of culture," *American Economic Review*, 106, 257–61.

OECD/ILO (2021): "Women at Work in G20 countries: Policy action since 2020," Technical report, OECD.

OLIVETTI, C., E. PATACCHINI, AND Y. ZENOU (2020): "Mothers, peers, and gender-role identity," *Journal of the European Economic Association*, 18, 266–301.

OLIVETTI, C. AND B. PETRONGOLO (2008): "Unequal pay or unequal employment? A cross-country analysis of gender gaps," *Journal of Labor Economics*, 26, 621–654.

PATNAIK, A., G. C. PAULEY, J. VENATOR, AND M. J. WISWALL (2023): "The Impacts of Same and Opposite Gender Alumni Speakers on Interest in Economics," Working Paper 30983, National Bureau of Economic Research.

POPE, D. G. AND J. R. SYDNOR (2010): "Geographic variation in the gender differences in test scores," *Journal of Economic Perspectives*, 24, 95–108.

PORTER, C. AND D. SERRA (2020): "Gender differences in the choice of major: The importance of female role models," *American Economic Journal: Applied Economics*, 12, 226–254.

RIVERA-GARRIDO, N. (2022): "Can education reduce traditional gender role attitudes?" *Economics of Education Review*, 89, 102261.

Rodríguez-Planas, N. and A. Sanz-de Galdeano (2019): "Intergenerational transmission of gender social norms and teenage smoking," *Social Science & Medicine*, 222, 122–132.

Rodríguez-Planas, N., A. Sanz-de Galdeano, and A. Terskaya (2022): "Gender norms in high school: Impacts on risky behaviors from adolescence to adulthood," *Journal of Economic Behavior & Organization*, 196, 429–456.

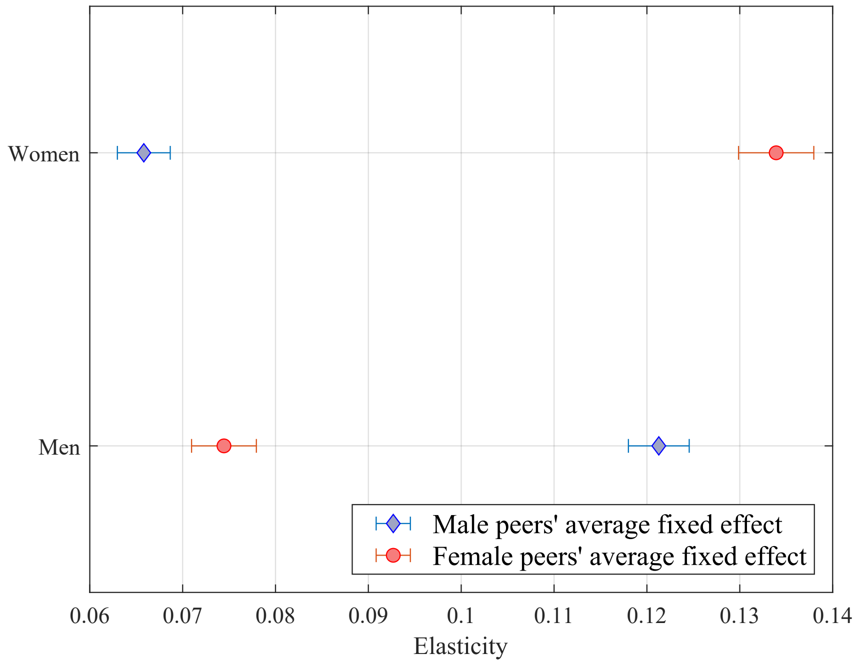
Rosaz, J., R. Slonim, and M. C. Villeval (2016): "Quitting and peer effects at work," *Labour Economics*, 39, 55–67.

Soetevent, A. R. and P. Kooreman (2007): "A discrete-choice model with social interactions: with an application to high school teen behavior," *Journal of Applied Econometrics*, 22, 599–624.

Van Veldhuizen, R., H. Oosterbeek, and J. Sonnemans (2018): "Peers at work: Evidence from the lab," *PloS one*, 13, e0192038.

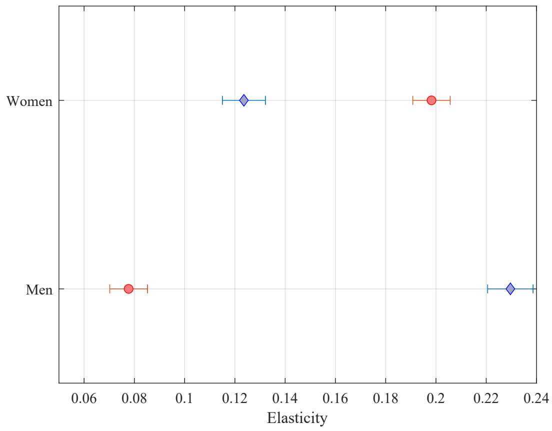
Tables and Figures

Figure 1: Same-Gender and Opposite-Gender Peer Effect on Wages

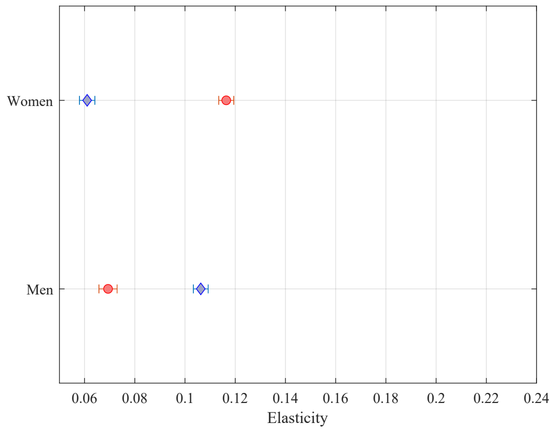


Notes: this figure reports the estimated effect of average same-gender and opposite-gender peers' average fixed effect on log wages (see equation 6) for men and women. All specifications control for age, age-squared, tenure, tenure-squared, worker fixed effect, occupation-establishment, occupation-year, and establishment-year fixed effects. Bootstrapped standard errors clustered at the establishment level are displayed. $N = 27,464,523$.

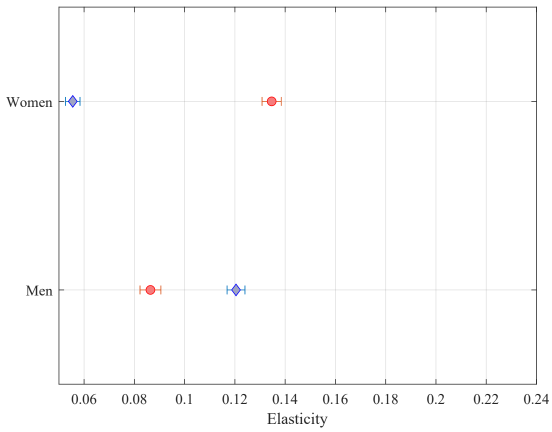
Figure 2: Same-Gender and Opposite-Gender Peer Effect on Wages by Establishment and Occupation Characteristics



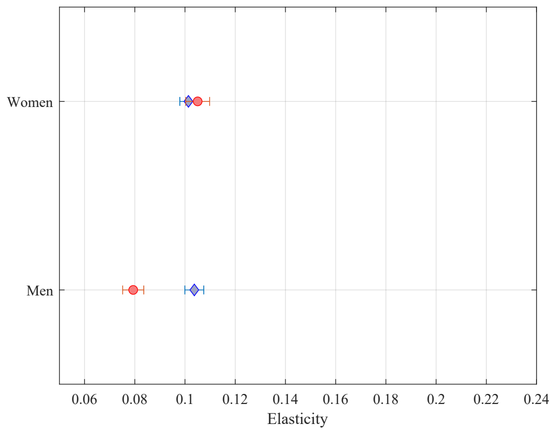
(a) Bottom 50% of skilled occupations



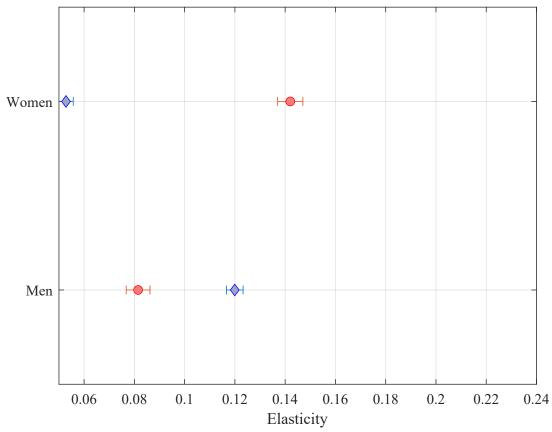
(b) Top 50% of skilled occupations



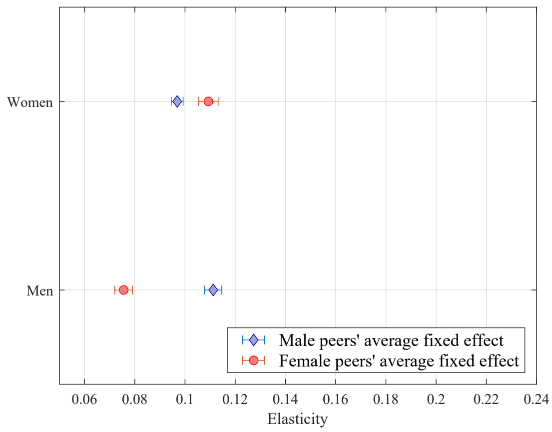
(c) Top quartile of gender wage gap



(d) Bottom quartile of gender wage gap



(e) Low share of women on top 25%



(f) High share of women on top 25%

Notes: this figure reports the estimated effect of average same-gender and opposite-gender peers' average fixed effect on individual log wages (see equation 6) by gender. All specifications control for age, age-squared, tenure, tenure-squared, worker fixed effect, occupation-establishment, occupation-year, and establishment fixed effects. Bootstrapped standard errors clustered at the firm level are displayed. The skill level of an occupation is defined according to the share of its tertiary-educated workers. The wage gender gap is computed at the establishment level after removing the effects of education, age, and year fixed effects. Low (high) share of women on top 25% is defined as 25% of firms with the lowest (highest) share of women on top 25% of firm's wage distribution. $N = 27,464,523$.

Table 1: Panel Structure of the Sample

Number of worker-year observations	27,464,523
Number of workers	7,583,748
Number of establishments	111,110
Number of occupations	569
Number of establishment-occupations	253,171
Number of establishment-years	504,542
Number of peer groups (establishment-occupation-years)	913,953
Average number of peer groups per establishment-year	1.811
Average number of occupations per worker	1.450
Average number of establishment per worker	1.595
Average number of occupations-establishment per worker	1.765
Average number of time periods per worker	3.621

Note: Statistics are based on RAIS data for the period 2003-2018. The overall sample has been constructed using the criteria outlined in Section 3.1.

Table 2: Summary Statistics by Gender

	Total	Men	Women
Panel A: Sample Description			
Real Monthly Wage (BRL 2005)	1923.435	2213.743	1624.811
Log of wage	7.104	7.220	6.986
Less than Primary Educ	0.085	0.100	0.070
Primary Educ.	0.123	0.139	0.108
Secondary Educ	0.525	0.500	0.551
Post-Secondary Educ.	0.234	0.227	0.241
Panel B: Peer Group Size			
Mean	30.050	30.718	29.396
Median	10.000	10.000	10.000
Panel C: Standard Deviation of Wages			
Raw Std. Dev. of log Real Wage	0.849	0.877	0.802
Standard Deviation of log Real Wage Residuals	0.563	0.585	0.540
Average Within Peer-Group Std. Dev. of log Real Wage Residuals	0.382	0.383	0.381
N	27,464,523	13,926,191	13,538,332

Note: Residuals used to calculate the standard deviation of log real wage residuals are derived from a log-wage regression that includes controls for time fixed effects, education, and quadratic terms for age and establishment tenure.

Table 3: Peer Quality

	Total	Males	Females
Standard deviation worker fixed effects	0.612	0.658	0.547
Standard deviation average male peers' fixed effects	0.524	0.550	0.490
Standard deviation average female peers' fixed effects	0.486	0.513	0.452
Standard deviation change of average male peers' fixed effects between t - 1 and t	0.176	0.174	0.178
Standard deviation change of average female peers' fixed effects between t - 1 and t	0.159	0.166	0.152
Standard deviation change of average male peers' fixed effects between t - 1 and t -Movers	0.287	0.302	0.271
Standard deviation change of average female peers fixed effects between t - 1 and t - Movers	0.320	0.321	0.319
Standard deviation change of average male peers' fixed effects between t - 1 and t -Stayers	0.099	0.104	0.095
Standard deviation change of average female peers' fixed effects between t - 1 and t - Stayers	0.107	0.105	0.109
Correlation worker fixed effects and average male peers' fixed effects	0.781	0.797	0.753
Correlation worker fixed effects and average female peers' fixed effects	0.772	0.767	0.779
Correlation occupation-time effects and average male peers' fixed effects	0.095	0.086	0.097
Correlation occupation-time effects and average female peers' fixed effects	0.082	0.062	0.099
Correlation establishment-time effects and average male peers' fixed effects	0.116	0.111	0.113
Correlation establishment-time effects and average female peers' fixed effects	0.097	0.081	0.109
Correlation occupation-establishment effects and average male peers' fixed effects	0.322	0.307	0.326
Correlation occupation-establishment effects and average female peers' fixed effects	0.300	0.269	0.328
N	27,464,523	13,926,191	13,538,332

Notes: Worker fixed effects are estimated using equation (6) and the algorithm described in Section 2.4.

Table 4: Gender Asymmetries in Peer Effects on Wages

The effect on	(1) Males	(2) Females
Male peers' average fixed effect	0.121	0.066
	(0.002)	(0.001)
Female peers' average fixed effect	0.074	0.134
	(0.002)	(0.002)
N	13,926,191	13,538,332
Worker FE	yes	yes
Occupation-year FE	yes	yes
Occupation-establishment FE	yes	yes
Establishment-year FE	yes	yes

Notes: the table reports the estimated effect of the same-gender and opposite-gender peers' average fixed effect on log wages by gender. All specifications control for quadratic forms of age and establishment tenure. Bootstrapped standard errors clustered at the establishment level are displayed in parentheses.

Table 5: Sensitivity Checks

The effect of...	(1) Female peers' average fixed effect	(2)	(3) Male peers' average fixed effect	(4)
	Males	Females	Males	Females
Panel A: Stayers				
Elasticity	0.083	0.137	0.123	0.071
	(0.002)	(0.002)	(0.002)	(0.002)
Panel B: Movers				
Elasticity	0.088	0.147	0.109	0.053
	(0.002)	(0.002)	(0.002)	(0.002)
Panel C: Occupation at 6-digit level				
Elasticity	0.092	0.142	0.126	0.079
	(0.001)	(0.001)	(0.001)	(0.001)
Panel D: Low-segregated Peer Groups				
Elasticity	0.091	0.133	0.107	0.068
	(0.002)	(0.002)	(0.002)	(0.002)
Panel E: With gender-specific fixed effects				
Elasticity	0.082	0.126	0.118	0.076
	(0.002)	(0.001)	(0.001)	(0.001)
Panel F: With peer-group-specific controls				
Elasticity	0.099	0.153	0.145	0.095
	(0.001)	(0.001)	(0.001)	(0.001)
Panel G: Peer group size 10-50				
Elasticity	0.077	0.144	0.130	0.068
	(0.002)	(0.002)	(0.002)	(0.002)
Panel H: Entire Brazil				
Elasticity	0.073	0.175	0.169	0.073
Panel I: Placebo Establishments				
Elasticity	−0.004	−0.000	0.002	0.001
	(0.000)	(0.000)	(0.000)	(0.000)
Panel J: Placebo Occupations				
Elasticity	−0.002	0.001	−0.000	−0.001
	(0.000)	(0.000)	(0.000)	(0.000)

Notes: The table reports the estimated effect of the same-gender and opposite-gender peers' average fixed effect on log wage by gender. Stayers are workers who did not change the occupation-firm. Movers are workers who change occupation-establishment. Low-segregated are peer groups with 40-60% of women. Placebo occupation is a specification where peer groups consist of workers from the same establishment but from randomly chosen occupations. Placebo establishments is a specification where peer groups are defined as workers from the same occupation but randomly chosen establishments. All specifications control for quadratic forms of age and firm tenure, worker fixed effects, occupation-year effects, establishment-year effects, and occupation-establishment fixed effects. Bootstrapped standard errors clustered at the establishment level are displayed in parentheses.

Table 6: Gender Asymmetries in Peer Effects by Establishment and Occupation Characteristics

The effect of...	(1)	(2)	(3)	(4)
	Female peers' average fixed effect		Male peers' average fixed effect	
	Males	Females	Males	Females
Panel A: Top 50% Skilled Occupation				
Elasticity	0.069	0.116	0.106	0.061
	(0.002)	(0.002)	(0.001)	(0.002)
Panel B: Bottom 50% Skilled Occupation				
Elasticity	0.078	0.198	0.230	0.124
	(0.004)	(0.004)	(0.005)	(0.004)
Panel C: Establishments at the Bottom Quartile of the Female-Male Wage Gap				
Elasticity	0.086	0.135	0.120	0.055
	(0.002)	(0.002)	(0.002)	(0.001)
Panel D: Establishments at the Top Quartile of the Female-Male Wage Gap				
Elasticity	0.079	0.105	0.104	0.101
	(0.002)	(0.002)	(0.002)	(0.002)
Panel E: Establishments at the Bottom Quartile of the Share of Women at the Top 25% of the Wage Dist.				
Elasticity	0.081	0.142	0.120	0.053
	(0.002)	(0.003)	(0.002)	(0.001)
Panel F: Establishments at the Top Quartile of the Share of Women at the Top 25% of the Wage Dist.				
Elasticity	0.076	0.109	0.111	0.097
	(0.002)	(0.002)	(0.002)	(0.001)

Notes: The table reports the estimated effect of the same-gender and opposite-gender peers' average fixed effect on log wages by gender and by establishment and occupation characteristics. All specifications control for quadratic forms of age and establishment tenure, worker fixed effects, occupation-year effects, establishment-year effects, and occupation-establishment fixed effects. Bootstrapped standard errors clustered at the establishment level are displayed in parentheses.

Appendix (For Online Publication)

Appendix A Theoretical Model Derivations

To guide our empirical estimation, this section develops a simple model of social interactions in the workplace that incorporates elements of the economics of identity Akerlof and Kranton (2000, 2002, 2005, 2008). In the model, peers' output affects individuals' actions through peer pressure. Given the features of our administrative data, described in Section 3, we assume in the model that individuals may identify themselves with two gender categories, men and women. When gender norms are egalitarian, the productivity of all coworkers affects an individual's wages in a similar fashion, regardless of their gender. In contrast, when gender norms are non egalitarian, the productivity of coworkers from the same social category may have a greater impact.

Production and effort functions. Worker i produces according to the following function:

$$f_{it} = \alpha_i + e_i + \epsilon_i \tag{A.1}$$

, where α_i is individual i's ability (or permanent component of productivity), e_i is individual i's effort, and ϵ_i is a random component of productivity independent of ability and effort. Ability is continuous and exogenously given such that $\alpha_i \in [0, A]$ and is distributed with probability density function $h(\alpha_i)$. Instead, individual effort is chosen by the individual to maximize utility. We assume that exerting effort is costly, following a quadratic cost function defined by:

$$C(e_i) = ce_i^2 \tag{A.2}$$

, where $c > 0$ is a scale parameter.

Peer pressure. Workers' utility depends on peer pressure stemming from coworkers, as in Kandel and Lazear (1992), Mas and Moretti (2009), and Cornelissen et al. (2017). If worker i deviates from her coworkers' production, peer pressure reduces her utility. This function can be parameterized to be increasing in the distance between a worker's output and the expected value of coworkers' output as follows:

$$P\Big(f_{it}, \mathbb{E}(f)\Big) = \delta\Big(\mathbb{E}(f) - f_{it}\Big)^2 \tag{A.3}$$

, where $\mathbb{E}(f)$ is the expected value of the production of workers and δ is a scale parameter that denotes how painful peer pressure is.

If social categories matter, individuals may feel different degrees of peer pressure depending on their gender and the gender of their peers. For instance, individuals may feel more social pressure if they deviate from the output of peers of the same gender than if they deviate from the output of peers of the opposite gender. In this case, the peer pressure function can be given by:

$$P\Big(f_{it}, \mathbb{E}(f_{c_{it}}), \mathbb{E}(f_{\notin c_{it}})\Big) = (\eta + \eta_s)(\mathbb{E}(f_{c_{it}}) - f_{it})^2 + \eta(\mathbb{E}(f_{\notin c_{it}}) - f_{it})^2 \tag{A.4}$$

, where $\mathbb{E}(f_{c_{it}})$ is the expected value of the production of workers who belong to the same social category as i and $\mathbb{E}(f_{\notin c_{it}})$ is the expected value of production of workers who belong to a different social category than i. There are two social categories, women and men, such that $c_i \in \{W, M\}$. For instance, if i is a woman ($c_i = W$), $\mathbb{E}(f_{c_{it}}) = \mathbb{E}(f_W)$ and $\mathbb{E}(f_{\notin c_{it}}) = \mathbb{E}(f_M)$.[17]

The parameter $\eta \geq 0$ in equation A.4 represents the pain of peer pressure that is independent of the social category of the worker. Instead, $\eta_s \geq 0$ represents potential asymmetries in the level of peer pressure. We assume that η_s depends on gender norms. If $\eta_s = 0$, gender norms are egalitarian, and all workers experience the same level of peer pressure from male and female peers, irrespective of their gender.

[17]See Appendix A for derivations of $\mathbb{E}(f_M)$ and $\mathbb{E}(f_W)$.

However, if $\eta_s > 0$, individuals experience more peer pressure from the same social category peers than from different social category peers. The higher is η_s, the less egalitarian gender norms are.

Optimal level of effort. Workers chose a level of effort that maximizes the difference between their earnings, the costs of providing effort, and the pain associated with peer pressure. The problem can be written as follows:

$$Max_{e_i}\mathbb{E}\Big[b_{c_i}f_{it} - ce_i^2 - P\Big(f_{it}, \mathbb{E}(f_{c_{it}}), \mathbb{E}(f_{\notin c_{it}})\Big)\Big] \tag{A.5}$$

, where b is the wage rate.

The first order condition of this problem is:

$$b - 2ce_i + 2(\eta + \eta_s)(\mathbb{E}(f_{c_{it}}) - f_{it}) + 2\eta(\mathbb{E}(f_{\notin c_i} - f_{it}) = 0 \tag{A.6}$$

Substituting $\mathbb{E}(f_{c_{it}}) = \mathbb{E}(\alpha_{c_i}) + \mathbb{E}(e_{c_i})$ and $\mathbb{E}(f_{\notin c_i}) = \mathbb{E}(\alpha_{\notin c_i}) + \mathbb{E}(e_{\notin c_i})$, we obtain:

$$e_i = \frac{\frac{b}{2} + (\eta + \eta_s)(\mathbb{E}(\alpha_{c_i}) + \mathbb{E}(e_{c_i})) - (2\eta + \eta_s)\alpha_i + \eta(\mathbb{E}(\alpha_{\notin c_i}) + \mathbb{E}(e_{\notin c_i}))}{c + 2\eta + \eta_s} \tag{A.7}$$

Integrating e_i across individuals of social category c_i, we obtain:

$$\mathbb{E}(e_{c_i}) = \frac{\frac{b}{2} + \eta(\mathbb{E}(e_{\notin c_i}) - \mathbb{E}(\alpha_{c_i})) + \eta(\mathbb{E}e_{\notin c_i})}{c + \eta} \tag{A.8}$$

Similarly:

$$\mathbb{E}(e_{\notin c_i}) = \frac{\frac{b}{2} + \eta(\mathbb{E}(\alpha_{c_i}) - \mathbb{E}(\alpha_{\notin c_i})) + \eta\mathbb{E}(e_{c_i})}{c + \eta} \tag{A.9}$$

Substituting (A.9) into (A.8), we obtain:

$$\mathbb{E}(e_{c_i}) = \frac{b}{2c} + \frac{c\eta(\mathbb{E}(\alpha_{c_i}) - \mathbb{E}(\alpha_{\notin c_i}))}{c(c + 2\eta)} \tag{A.10}$$

$$\mathbb{E}(e_{\notin c_i}) = \frac{b}{2c} + \frac{c\eta(\mathbb{E}(\alpha_{\notin c_i}) - \mathbb{E}(\alpha_{c_i}))}{c(c+2\eta)} \tag{A.11}$$

Substituting (A.10) and (A.11) into (A.7), effort function can be written as:

$$e_i = \tau\alpha_i + b(\rho_s + \rho_o) + \theta_s\mathbb{E}(\alpha_{c_i}) + \theta_o\mathbb{E}(\alpha_{\notin c_i}) \tag{A.12}$$

, where $\tau = \frac{-(2\eta+\eta_s)}{c+2\eta+\eta_s}$, $\rho_s = \frac{\eta+c}{2c(c+2\eta)}$, $\rho_o = \frac{\eta}{2c(c+2\eta)}$, $\theta_o = \frac{\eta}{c+2\eta}$, and $\theta_s = \frac{(\eta+\eta_s)(c+\eta)+\eta^2}{(c+2\eta)(c+2\eta+\eta_s)}$.

Because $\theta_s > 0$ and $\theta_o > 0$, the output of peers enters positively into the effort function. Moreover, note that

$$\theta_s - \theta_o = \frac{c\eta_s}{(c+2\eta)(c+2\eta+\eta_s)} \tag{A.13}$$

, which implies that the effect of same social category peers on effort is greater than the effect of opposite social category peers if and only if $\eta_s > 0$. If the social pressure function is symmetric ($\eta_s = 0$), same and opposite social categories have identical impacts on workers' effort, i.e., $\theta_o = \theta_s$. Note that $\theta_s - \theta_o$ is monotonically increasing in η_s,[18] i.e., the less egalitarian gender norms are, the larger will be the gap in the response of workers to peers of the same gender with respect to their response to peers of the opposite gender.

Introducing (A.12) into (A.1), we obtain an expression for workers' earnings:

$$w_i = bf_{it} = b[(1+\tau)\alpha_i + b(\rho_s + \rho_o) + \theta_s\mathbb{E}(\alpha_{c_i}) + \theta_o\mathbb{E}(\alpha_{\notin c_i}) + \epsilon_i]. \tag{A.14}$$

Individual earnings are positively affected by workers' ability and wages, but they also depend on the ability of peers of the same and opposite social categories. The objective of our empirical analysis is to estimate θ_o and θ_s, which as discussed above are informative about η_s. Note however that we cannot directly bring equation (A.12) to the data because α_i, α_{c_i}, and $\alpha_{\notin c_i}$ are not observable. We instead estimate them using worker fixed effects, as described in Section 2. But as equation (A.12) highlights,

[18]If $\eta_s > 0$, then $\frac{d(\theta_s - \theta_o)}{d\eta_s} = \frac{c}{(c+\eta_s+2\eta)^2} > 0$.

the estimate of a worker fixed effect is a combination of individual ability and other determinants of earnings summarized by $b(1+\tau)\alpha_i$.

To better understand the impact of approximating ability by workers' fixed effects on the interpretation of the results, let us define workers' fixed effect as $\tilde{\alpha}_i = b(1+\tau)\alpha_i$. This fixed effect can be interpreted as a monetary representation of individual ability. Averaging across individuals in c_i, we obtain $\mathbb{E}(\tilde{\alpha}_{c_i}) = b(1+\tau)\mathbb{E}(\alpha_{c_i})$. Similarly, $\mathbb{E}(\tilde{\alpha}_{\notin c_i}) = b(1+\tau)\mathbb{E}(\alpha_{\notin c_i})$. Thus, we can rewrite (A.14) in terms of $\tilde{\alpha}_i$:

$$w_i = \kappa + \tilde{\alpha}_i + \tilde{\theta}_s \mathbb{E}(\tilde{\alpha}_{c_i}) + \tilde{\theta}_o \mathbb{E}(\tilde{\alpha}_{\notin c_i}) + v_i \tag{A.15}$$

, where $\kappa = b(\rho_s + \rho_o)$, $v_i = b\epsilon_i$, $\tilde{\theta}_s = \frac{\theta_s}{1+\tau}$, $\tilde{\theta}_o = \frac{\theta_o}{1+\tau}$.

Note that $\tilde{\theta}_s - \tilde{\theta}_o = \frac{\eta_s}{c+2\eta}$ given the results from equation (A.13) and the definition of τ. Therefore, $\tilde{\theta}_s - \tilde{\theta}_o > 0$ if and only if $\eta_s > 0$. Moreover, it is strictly increasing on η_s.

Appendix B Proof of Theorem 1

Let us start by specifying the first-order condition for a_i for women and men separately. $\mathbf{M}_{nt}$ and $\mathbf{F}_{nt}$ denote the sets of male and female workers in peer group n in period t, respectively.

FOC for women:

$$\begin{aligned} &\sum_{t=1}^{T} \left(w_{itp} - a_i - \frac{\alpha_1}{F_{tn}-1} \sum_{f \in \mathbf{F}_{nt \sim i}} a_f - \frac{\beta_1}{M_{tn}} \sum_{m \in \mathbf{M}_{nt}} a_m \right) + \\ &\sum_{t=1}^{T} \sum_{f \in \mathbf{F}_{nt \sim i}} \frac{\alpha_1}{F_{tn}-1} \left(w_{ftn} - a_f - \frac{\alpha_1}{F_{tn}-1} \Big(\sum_{l \in \mathbf{F}_{nt \sim i \sim f}} a_l + ai \Big) - \frac{\beta_1}{M_{tn}} \sum_{m \in \mathbf{M}_{nt}} a_m \right) + \\ &\sum_{t=1}^{T} \sum_{m \in \mathbf{M}_{nt}} \frac{\alpha_2}{F_{tn}} \left(w_{mtn} - a_m - \frac{\alpha_2}{F_{tn}} \Big(\sum_{k \in \mathbf{F}_{nt \sim i}} a_k + ai \Big) - \frac{\beta_2}{M_{tn}-1} \sum_{j \in \mathbf{M}_{nt} \sim m} a_j \right) = 0 \end{aligned} \tag{B.1}$$

FOC for men:

$$\sum_{t=1}^{T}\left(w_{itp}-a_i-\frac{\alpha_2}{F_{tn}}\sum_{f\in\mathbf{F}_{nt}}a_f-\frac{\beta_2}{M_{tn}}\sum_{m\in\mathbf{M}_{nt\sim i}}a_m\right)+$$
$$\sum_{t=1}^{T}\sum_{f\in\mathbf{F}_{nt}}\frac{\beta_1}{M_{tn}}\left(w_{ftn}-a_f-\frac{\alpha_1}{F_{tn}-1}\sum_{l\in\mathbf{F}_{nt\sim f}}a_l-\frac{\beta_1}{M_{tn}}(\sum_{m\in\mathbf{M}_{nt\sim i}}a_m+a_i)\right)+ \tag{B.2}$$
$$\sum_{t=1}^{T}\sum_{m\in\mathbf{M}_{nt\sim i}}\frac{\beta_2}{M_{tn}-1}\left(w_{mtn}-a_m-\frac{\alpha_2}{F_{tn}}\sum_{k\in\mathbf{F}_{nt}}a_k-\frac{\beta_2}{M_{tn}-1}(\sum_{j\in\mathbf{M}_{nt}\sim m\sim i}a_j+a_i)\right)=0$$

Solving for a_i and collecting terms we have for women:

$$a_i=\sum_{t=1}^{T}\left[\left(w_{itp}-\frac{\alpha_1}{F_{tn}-1}\sum_{f\in\mathbf{F}_{nt\sim i}}a_f-\frac{\beta_1}{M_{tn}}\sum_{m\in\mathbf{M}_{nt}}a_m\right)+\right.$$
$$\sum_{f\in\mathbf{F}_{nt\sim i}}\frac{\alpha_1}{F_{tn}-1}\left(w_{ftn}-a_f-\frac{\alpha_1}{F_{tn}-1}\sum_{l\in\mathbf{F}_{nt\sim i\sim f}}a_l-\frac{\beta_1}{M_{tn}}\sum_{m\in\mathbf{M}_{nt}}a_m\right)+ \tag{B.3}$$
$$\left.\sum_{m\in\mathbf{M}_{nt}}\frac{\alpha_2}{F_{tn}}\left(w_{mtn}-a_m-\frac{\alpha_2}{F_{tn}}\sum_{k\in\mathbf{F}_{nt\sim i}}a_k-\frac{\beta_2}{M_{tn}-1}\sum_{j\in\mathbf{M}_{nt}\sim m}a_j\right)\right]/Den_{fn}$$

where $Den_{fn}=\sum_{t=1}^{T}\left(1+\frac{\alpha_1^2}{F_{nt}-1}+\frac{M_{nt}\alpha_2^2}{F_{nt}^2}\right)$

And for men:

$$a_i=\sum_{t=1}^{T}\left[\left(w_{itp}-\frac{\alpha_2}{F_{tn}}\sum_{f\in\mathbf{F}_{nt}}a_f-\frac{\beta_2}{M_{tn}}\sum_{m\in\mathbf{M}_{nt\sim i}}a_m\right)+\right.$$
$$\sum_{f\in\mathbf{F}_{nt}}\frac{\beta_1}{M_{tn}}\left(w_{ftn}-a_f-\frac{\alpha_1}{F_{tn}-1}\sum_{l\in\mathbf{F}_{nt\sim f}}a_l-\frac{\beta_1}{M_{tn}}\sum_{m\in\mathbf{M}_{nt\sim i}}a_m\right)+ \tag{B.4}$$
$$\left.\sum_{t=1}^{T}\sum_{m\in\mathbf{M}_{nt\sim i}}\frac{\beta_2}{M_{tn}-1}\left(w_{mtn}-a_m-\frac{\alpha_2}{F_{tn}}\sum_{k\in\mathbf{F}_{nt}}a_k-\frac{\beta_2}{M_{tn}-1}\sum_{j\in\mathbf{M}_{nt}\sim m\sim i}a_j\right)\right]/Den_{mn}$$

where $Den_{mn}=\sum_{t=1}^{T}\left(1+\frac{\beta_1^2F_{nt}}{M_{nt}^2}+\frac{\beta_2^2}{M_{nt}^2}\right)$

The iterative method starts by making a first guess about vector a, using this first guess we then generate OLS estimates of the parameters of the model. Then these estimates are plugged into the RHS of the FOC and a_i are updated accordingly. Following Arcidiacono et al. (2012), let us call the first guess a and the next guess a'. We next show that the mapping function $f: a \rightarrow a'$ provided by equations (B.3) and (B.4) is a contraction mapping. That is, $d(f(a), f(a')) < \beta d(a, a')$ for some $\beta < 1$, where d is a valid distance function. We use an Euclidian distance function for d to show the conditions under which f is a contraction mapping for some $\beta < 1$. Let us define $\tilde{a} = a - a'$. Summing up (B.3) and (B.4) into one vector, we can write the condition as follows:

$$\left[\sum_{i=1}^{N_f}\left(-\sum_{t=1}^{T}(\gamma_{1nt}\sum_{f\in\mathbf{F}_{nt\sim i}}\tilde{a}_f+\delta_{nt}\sum_{m\in\mathbf{M}_{nt}}\tilde{a}_m)/Den_{fn}\right)^2+\right.$$
$$\left.\sum_{i=1}^{N_m}\left(-\sum_{t=1}^{T}(\delta_{nt}\sum_{f\in\mathbf{F}_{nt}}\tilde{a}_f+\gamma_{2nt}\sum_{m\in\mathbf{M}_{nt\sin i}}\tilde{a}_m)/Den_{mn}\right)^2\right]^{1/2} < \tag{B.5}$$
$$\beta\left(\sum_{i=1}^{N_f}\tilde{a}_i{}^2+\sum_{j=1}^{N_m}\tilde{a}_j{}^2\right)^{1/2}$$

where N_f refers to the total female population and N_m to the total male population, and:

$$\gamma_{1nt} = \frac{2\alpha_1}{F_{nt}-1}+\frac{\alpha_1^2(F_{nt}-2)}{(F_{nt}-1)^2}+\frac{\alpha_2^2M_{nt}}{F_{nt}^2} \tag{B.6}$$

$$\gamma_{2nt} = \frac{2\beta_2}{M_{nt}-1}+\frac{\beta_2^2(M_{nt}-2)}{(M_{nt}-1)^2}+\frac{\beta_1^2F_{nt}}{M_{nt}^2} \tag{B.7}$$

$$\delta_{nt} = \frac{\beta_1}{M_{nt}(1+\alpha_1}+\frac{\alpha_2}{F_{nt}}(1+\beta_2) \tag{B.8}$$

Next we simplify this inequality using Cauchy-Schwarz Inequality (CSI). Note that this transformation increases the LHS making it less likely that the inequality is satisfied.

$$\left[\sum_{i=1}^{N_f}\sum_{t=1}^{T}T\left(\gamma_{1nt}\sum_{f\in\mathbf{F}_{nt\sim i}}\tilde{a}_f+\delta_{nt}\sum_{m\in\mathbf{M}_{nt}}\tilde{a}_m\right)^2/Den_{fn}^2+\right.$$
$$\left.\sum_{i=1}^{N_m}\sum_{t=1}^{T}T\left(\delta_{nt}\sum_{f\in\mathbf{F}_{nt}}\tilde{a}_f+\gamma_{2nt}\sum_{m\in\mathbf{M}_{nt\sin i}}\tilde{a}_m\right)^2/Den_{mn}^2\right]^{1/2} \tag{B.9}$$
$$<\beta\left(\sum_{i=1}^{N_f}\tilde{a}_i{}^2+\sum_{j=1}^{N_m}\tilde{a}_j{}^2\right)^{1/2}$$

Applying CSI again yields:

$$
\begin{aligned}
&\left[\sum_{i=1}^{N_f}\sum_{t=1}^{T} T\left(\frac{\gamma_{1nt}^2+\delta_{nt}^2}{Den_{fn}^2}\right)\left(\Big(\sum_{f\in\mathbf{F}_{nt\sim i}}\tilde{a}_f\Big)^2+\Big(\sum_{m\in\mathbf{M}_{nt}}\tilde{a}_m\Big)^2\right)+\right.\\
&\left.\sum_{i=1}^{N_m}\sum_{t=1}^{T} T\left(\frac{\gamma_{2nt}^2+\delta_{nt}^2}{Den_{mn}^2}\right)\left(\Big(\sum_{f\in\mathbf{F}_{nt}}\tilde{a}_f\Big)^2+\Big(\sum_{m\in\mathbf{M}_{nt\sim i}}\tilde{a}_m\Big)^2\right)\right]^{1/2}\\
&\qquad < \beta\left(\sum_{i=1}^{N_f}\tilde{a}_i{}^2+\sum_{j=1}^{N_m}\tilde{a}_j{}^2\right)^{1/2}
\end{aligned}
\tag{B.10}
$$

Expanding the square of the sum of $\tilde{a}_i$'s and applying CSI:

$$
\begin{aligned}
&\left[\sum_{i=1}^{N_f}\sum_{t=1}^{T} T\left(\frac{\gamma_{1nt}^2+\delta_{nt}^2}{Den_{fn}^2}\right)\left((F_{nt}-1)\sum_{f\in\mathbf{F}_{nt\sim i}}\tilde{a}_f^2+M_{nt}\sum_{m\in\mathbf{M}_{nt}}\tilde{a}_m^2\right)+\right.\\
&\left.\sum_{i=1}^{N_m}\sum_{t=1}^{T} T\left(\frac{\gamma_{2nt}^2+\delta_{nt}^2}{Den_{mn}^2}\right)\left(F_{nt}\sum_{f\in\mathbf{F}_{nt}}\tilde{a}_f^2+(M_{nt}-1)\sum_{m\in\mathbf{M}_{nt\sim i}}\tilde{a}_m^2\right)\right]^{1/2}\\
&\qquad < \beta\left(\sum_{i=1}^{N_f}\tilde{a}_i{}^2+\sum_{j=1}^{N_m}\tilde{a}_j{}^2\right)^{1/2}
\end{aligned}
\tag{B.11}
$$

Since the terms γ_{2nt}, γ_{2nt}, δ_{nt}, Den_{fn}, Den_{mn}, F_{nt} and M_{nt} reflect differences in peer group sizes experienced by individual i over time, all the terms on the left hand side will have different multipliers. To address this issue, we substitute all the terms in the numerator (γ_{2nt}, γ_{2nt}, δ_{nt}, F_{nt} and M_{nt}) by their maximum values (denoted by γ_1, γ_2, δ, M, and F) and all the terms in denominator (Den_{mn}, Den_{fn}) by their minimum values (denoted by Den_m, Den_f). Note that this transformation is valid since it will strictly increase the LHS making it less likely that the inequality is satisfied.

This transformation leaves us with:

$$
\begin{aligned}
&\left[\sum_{i=1}^{N_f} T^2\left(\frac{\gamma_1^2+\delta^2}{Den_f^2}(F-1)^2+\frac{\gamma_2^2+\delta^2}{Den_m^2}FM\right)\tilde{a}_i^2+\sum_{i=1}^{N_m} T^2\left(\frac{\gamma_2^2+\delta^2}{Den_m^2}(M-1)^2+\frac{\gamma_1^2+\delta^2}{Den_f^2}FM\right)\tilde{a}_i^2\right]^{1/2}\\
&\qquad < \beta\left(\sum_{i=1}^{N_f}\tilde{a}_i{}^2+\sum_{j=1}^{N_m}\tilde{a}_j{}^2\right)^{1/2}
\end{aligned}
\tag{B.12}
$$

Finally, replacing F and M by their maximum value denoted by G and replacing (G-1) by G we arrive to the common multiplier:

$$T\left[\left(\frac{\gamma_1^2+\delta^2}{Den_f^2}G^2+\frac{\gamma_2^2+\delta^2}{Den_m^2}G^2\right)\right]^{1/2}\left(\sum_{i=1}^{N_f}\tilde{a}_i{}^2+\sum_{j=1}^{N_m}\tilde{a}_j{}^2\right)^{1/2} < \beta\left(\sum_{i=1}^{N_f}\tilde{a}_i{}^2+\sum_{j=1}^{N_m}\tilde{a}_j{}^2\right)^{1/2} \tag{B.13}$$

Now we need to show for which parameter values
$T\left[\left(\frac{\gamma_1^2+\delta^2}{Den_f^2}G^2+\frac{\gamma_2^2+\delta^2}{Den_m^2}G^2\right)\right]^{1/2} < 1$

Next it can be shown that $\frac{T}{Den_{fn}}$ and $\frac{T}{Den_{mn}}$ are always lower or equal than 1. Therefore we need to show that $(G\gamma_1)^2 + 2(G\delta)^2 + (G\gamma_2)^2 < 1$. Substituting G in equations (B.6), (B.7), and (B.8), we obtain the following inequality:

$$\begin{gathered}\left(2\alpha_1+\alpha_1^2(G-2)G/(G-1)^2+\alpha_2^2\right)^2+\\ \left(2\beta_2+\beta_2^2(G-2)G/(G-1)^2+\beta_1^2\right)^2+\\ 2\left(\beta_1(1+\alpha_1)+\alpha_2(1+\beta_2)\right)^2<1\end{gathered} \tag{B.14}$$

Now we replace $(G-2)G/(G-1)^2$ by $G/(G-1)$, which increases the LHS making it less likely to hold. Since the maximum value of $G/(G-1)$ for $G \geq 2$ is 2, we can replace $(G-2)G/(G-1)^2$ by 2. We end up with the following condition for the parameters:

$$\left(2\alpha_1+2\alpha_1^2+\alpha_2^2\right)^2+\left(2\beta_2+2\beta_2^2+\beta_1^2\right)^2+2\left(\beta_1(1+\alpha_1)+\alpha_2(1+\beta_2)\right)^2<1 \tag{B.15}$$

It can be shown that the inequality will be satisfied when all the coefficients are below 0.2.[19]

Appendix C Monte Carlo Simulations

In order to assess the magnitude of the bias associated with (i) peer group specific shocks, and (ii) serial correlation in the individual error term, we conduct Monte-

[19] This can be shown assuming that $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = x$ and solving the inequality for x.

Carlo simulations. We simulate the dependent variable as follows:

- Predict log wages in our estimation sample using the coefficients of the control variables and the fixed effects obtained when estimating our baseline model. We consider several scenarios regarding the magnitude of peer effects:

  1. Peer effects are equal to zero.
  2. Same-gender and opposite-gender peer effects are both equal to 0.05.
  3. The effect of same-gender peers is equal to 0.1 and the effect of opposite-gender peers is equal to 0.05.
  4. Same-gender and opposite-gender peer effects are similar to our baseline estimates: the effect of same-gender peers is equal to 0.15 and the effect of opposite-gender peers is equal to 0.07.

- Simulate peer-group specific shocks as normally distributed errors composed of an idiosyncratic component and a peer-group specific component. We consider the following scenarios:

  1. No peer-group specific component.
  2. The variance of the peer-group shock as a share of the total error variance is equal to 0.0267, which is the R^2 of the regression of the residuals from our main specification on peer-group fixed effects.
  3. The variance of the peer-group shock as a share of the total variance is 0.06, which is the value considered in Cornelissen et al. (2017).

- Simulations of serially correlated errors. We add a normally distributed error term with variance equal to the estimated error variance from the baseline model. We also assume first-order serial correlation in the error term with an autocorrelation coefficient equal to 0.34, which is the value obtained when regressing the residuals from our baseline model on its lagged value.

- For each type of error we run 5 simulations. Therefore, we have a total of 80 simulated dependent variables (4 peer effect coefficients $\times$ 4 errors $\times$ 5 simulations).

- We then estimate our main specification for each simulated dependent variable and compute average peer effect coefficients over 5 simulations.

Table C.1: Monte-Carlo Simulations

The effect of...	(1) Female peers' average fixed effect: Males	(2) Female peers' average fixed effect: Females	(3) Male peers' average fixed effect: Males	(4) Male peers' average fixed effect: Females
Panel A				
True effect	0.000	0.000	0.000	0.000
i.i.d	0.001	0.000	-0.000	0.000
Peer group shock (3%)	0.007	0.006	0.005	0.006
Peer group shock (6%)	0.014	0.012	0.010	0.012
Serial correlation ($\rho = 0.34$)	0.000	-0.001	-0.000	0.000
Panel B				
True effect	0.050	0.050	0.050	0.050
i.i.d	0.049	0.049	0.049	0.049
Peer group shock (3%)	0.058	0.056	0.055	0.056
Peer group shock (6%)	0.066	0.062	0.059	0.063
Serial correlation ($\rho = 0.34$)	0.048	0.047	0.048	0.049
Panel C				
True effect	0.050	0.100	0.100	0.050
i.i.d	0.050	0.098	0.098	0.050
Peer group shock (3%)	0.058	0.106	0.104	0.057
Peer group shock (6%)	0.066	0.112	0.109	0.063
Serial correlation ($\rho = 0.34$)	0.048	0.096	0.096	0.049
Panel D				
True effect	0.070	0.150	0.150	0.070
i.i.d	0.069	0.146	0.147	0.069
Peer group shock (3%)	0.078	0.155	0.153	0.077
Peer group shock (6%)	0.086	0.161	0.159	0.084
Serial correlation ($\rho = 0.34$)	0.068	0.144	0.145	0.069

Notes: The table presents the results of Monte Carlo simulations. The data generation process (DGP) is described in Appendix C. The row labeled "True effect" shows the peer effects assumed in the DGP. The row labeled "i.i.d." shows the simulation results when the errors are assumed to be independently and identically normally distributed and have a variance equal to the variance of the residuals from our main specification given by equation (6). The rows labeled "Peer group shock (3%)" and "Peer group shock (6%)" show the simulation results when peer group-specific shocks constitute 3% and 6% of the total error variance, respectively. The row labeled "Serial correlation ($\rho = 0.34$)" shows the simulation results when the errors are generated with first-order serial correlation and an autocorrelation coefficient equal to 0.34. Each coefficient is computed as the average coefficient obtained from 5 simulations.