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

WHAT LIES BEHIND THE IMMIGRANT WAGE GAP:

THE ROLE OF LABOR MARKET SORTING AND PEER EFFECTS

JOÃO PEDRO BONITO GOMES

Work Project carried out under the supervision of:

Professor Pedro Portugal

09-01-2025

WHAT LIES BEHIND THE IMMIGRANT WAGE GAP:
THE ROLE OF LABOR MARKET SORTING AND PEER EFFECTS*

JOÃO BONITO GOMES†

January 9, 2025

ABSTRACT

In this study, we analyze the impact of immigrant and native peers' ability on immigrant and native workers' wages, using rich matched employer-employee data on the Portuguese labor market. We employ an estimation strategy that circumvents problems of endogeneity and show that, when both worker and peer are natives or immigrants, a one standard deviation increase in the average quality of peers increases wages by 2.21% or 1.18%, respectively. We also find that these peer effects account for 4.47% of the immigrant wage gap, while the establishment/occupation/year component explains 52.30%, and the worker component explains 43.23% of the wage differential.

JEL CODES: J15, J24, J31, J61, J71.

KEYWORDS: Immigrants, Wage gap, Labor market sorting, Peser effects, Gelbach decomposition, Matched employer-employee data, High-dimensional fixed effects model.

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

∗. I would like to express my deep gratitude to Professor Pedro Portugal for all the guidance during the development of this project, whose unique empirical knowledge and assistance made it possible. I would also like to extend my thanks to Daniel Castelo for all the help throughout the last months, and to João and Maria for the support and friendship. Aos meus pais e à minha irmã, por terem sempre acreditado nos meus sonhos, mas especialmente por terem sempre acreditado no poder da educação para mudar vidas.

†. Nova School of Business and Economics, Universidade Nova de Lisboa. Email: joao.b.gomes@novasbe.pt.

I Introduction

Portugal has historically been a country characterized by emigration, with large numbers of its citizens seeking opportunities abroad. However, since the end of the Portuguese Colonial War in 1974, the number of immigrants in Portugal steadily increased over the years. This rapid growth of the immigrant population in Portugal raises important questions about their economic performance and impact in the labor market. In this way, this study aims to assess what accounts (and how much) for the difference between the wages of natives and immigrants.

Differences in peer quality can potentially play an important role in explaining the immigrant wage gap by shaping the learning environment, opportunities for assimilation, and transfer of skills in the workplace.[1] Therefore, besides accounting for differences in workers' ability and their allocation to different establishments and occupations over time, this study also considers the impact of peers' quality on wages and their influence in the immigrant wage gap.

The core contribution of this study consists in incorporating the methodology of Arcidiacono et al. (2012) into the canonical model of Abowd, Kramarz, and Margolis (1999) (AKM Model), and then extend it by allowing for both heterogeneity in the response to peers and heterogeneity in the influence of peers based on nationality, i.e., the impact of immigrant (native) peers' ability on native (immigrant) workers' wages. We then use the method developed by Gelbach (2016) to assess the contribution of each component in explaining this wage differential. To the best of our knowledge, this is the first application of such an estimation strategy, enabling us to assess the impact of peers' quality on wages, while distinguishing between workers and peers based on their immigrant or native status.

Our results provide evidence that the impact of peers' average ability is greater when both the worker and peer belong to the same nationality-specific group (either native or immigrant)

1. In this study, "peer" and "coworker" are terms used synonymously.

compared to when they differ in nationality status. We also observe that these human capital spillovers have a modest impact on the immigrant wage gap (4.47%), when compared to the worker component (43.23%) and the establishment/occupation/year component (52.30%).

This study is structured as follows. Section II presents the main literature on the immigrant wage gap and potential factors over the years. Section III describes the data set and concepts used, and Section IV provides a brief exploratory analysis of the key characteristics distinguishing immigrants from natives. Section V outlines the estimation strategy, and Section VI presents the empirical results obtained. Finally, Section VII concludes.

II LITERATURE REVIEW

While the field of labor economics that studies wage differentials between distinct groups (e.g., by gender or ethnicity) usually resorts to traditional discrimination theories (see BECKER 1957, PHELPS 1972, and ARROW 1971, e.g.), the immigrant wage gap is often analyzed through the perspective of assimilation pioneered by CHISWICK (1978). Its objective is to understand how the immigrant wage gap evolves over time and which factors are most important in this process. Several experiments were conducted to investigate which aspects are involved in the economic assimilation mechanism. These studies focus on dimensions such as the country where an individual's human capital was acquired (FRIEDBERG 2000), the fluency in the local language (CHISWICK and MILLER 2015; AYDEMIR and SKUTERUD 2005), or where the worker's training and experience were obtained (COHEN-GOLDNER and ECKSTEIN 2002).

However, the economic literature addressing the immigrant wage gap in the Portuguese labor market is relatively scarce. The work by CARNEIRO, FORTUNA, and VAREJÃO (2012) suggested that immigrants are often offered lower wages due to the lack of match-specific human capital and occupational downgrading, with these two factors explaining two-thirds of the gap. Another study by CABRAL and DUARTE (2013), using the Gelbach decomposition method, obtained that the returns from education are lower for immigrants than for natives at all levels of education, and that foreign work experience is less valuable than experience acquired in Portugal, which aligns with the conclusions on imperfect transferability of human capital across borders presented by FRIEDBERG (2000). Using the AKM Model, thus controlling for both individual and firm fixed effects, DE MATOS (2016) examined the role of job mobility, and concluded that around 30% of the wage convergence between immigrants and natives can be attributed to immigrants moving to firms offering higher wage premiums.

Regarding on-the-job interactions, differences in peer quality can play an important role in explaining the immigrant wage gap by shaping the learning environment, opportunities for assimilation, and transfer of skills in the workplace. Although peer effects have already been theorized several decades ago (see MARSHALL 1890 and LUCAS 1988, e.g.), the empirical literature is scant. Some of theses studies were conducted in environments with specific characteristics or considered only one firm or occupation. For example, FALK and ICHINO (2006) resorted to a controlled field experiment to conclude that workers' productivity increased when they worked alongside more productive co-workers. Similarly, MAS and MORETTI (2009) used information from a large supermarket chain and found positive peer effects arising from social pressure.[2]

More recently, empirical research started to use larger and more representative data sets to assess peer effects. A selection of this research focus on the contemporaneous peer effects, i.e., the effects from peers within the same time period. For example, CORNELISSEN, DUSTMANN, and SCHÖNBERG (2017) found only small peer effects, using data from Munich,

2. Other papers focusing one a single occupation include, for example, the work by GURYAN, KROFT, and NOTOWIDIGDO (2009) that tested for peer effects among golf players in professional tournaments, but found no evidence of these effects, and the work by LINDQUIST, SAUERMANN, and ZENOU (2015) that, using data from call center workers, found evidence of strong network effects in peers' productivity. In related studies, KAUR, KREMER, and MULLAINATHAN (2010) exploited data from an Indian company and observed productivity spillovers among workers, while CHAN, LI, and PIERCE (2014) observed data from a department store sales in China.

Germany. Battisti (2017), using data from Veneto, Italy, and Portugal et al. (2022), using data from Portugal, identified that a 10% increase in peer quality was associated to a 3.6% and 2.0% wage premium, respectively.[3] On the other hand, Jarosch, Oberfield, and Rossi-Hansberg (2021) and Hong and Lattanzio (2022), also using data from Germany and Italy, respectively, analyzed the dynamics of peer effects, i.e., how the magnitude of these effects evolve over time. Both papers found positive peer effects, although, while Hong and Lattanzio (2022) argued that they erode over time, Jarosch, Oberfield, and Rossi-Hansberg (2021) claimed that they actually increase. Several of these studies performed heterogeneity analyzes to try to assess the relative importance of different characteristics that could be driving the results. In this regard, the findings suggested that there were stronger peer effects for younger and low-ability workers (Cornelissen, Dustmann, and Schönberg 2017; Hong and Lattanzio 2022). Finally, literature on the relation between spillover effects and wage differentials is still very scarce, with Battisti (2017) noting that 12% of the gender wage gap and 10% of the immigrant wage gap can be explained by differences in peer characteristics.

III Data

III.A Source

This study uses data from *Quadros de Pessoal* (*QP*), a longitudinal data set with information collected by the Portuguese Ministry of Labor, that matches workers, establishments, and firms based in Portugal. The administrative data is collected annually in a reference month (October) through a mandatory employment survey for all firms in Portugal with at least one wage earner, with the exception of civil servants, self-employed individuals, and household employees. It

3. Cornelissen, Dustmann, and Schönberg (2017) estimated an increase of 1.5% and 0.1% in wages in the most similar regressions to the ones defined by Battisti (2017) and Portugal et al. (2022), respectively.

presents unique advantages that go along with the objective of this study. First, it comprises detailed data on several characteristics of the worker (e.g, age, gender, nationality, education, occupation, earnings, hours worked, tenure, type of contract, etc.) and of the firm (e.g., location, economic activity, sales, etc.). Second, information on worker's earnings is also very complete, including the base wage, regular and non-regular benefits, and overtime payment. Furthermore, the identification number of workers, establishments, and firms allows us to trace them throughout the years in the sample, creating the conditions to apply high-dimensional fixed effects regression models.

III.B Sample Selection

Although this survey data started to be collected in 1985, it was only in 2000 that *QP* started to include information regarding the worker's nationality. However, since there is no data for 2001, we only use data referring the period between 2002 and 2022. We also restrict the sample to full-time workers between the age of 18 and 65, working at least 120 hours per month, reporting a job tenure between 0 and 50 years, and that were receiving a base wage of at least 80% of the minimum legal wage in Portugal.[4] We also exclude workers in sectors and industries that normally provide services to other firms through outsourcing (e.g., cleaning, security, or financial services), so that we guarantee that workers effectively share the same physical workplace when forming peer groups. Finally, any individual with missing values on the variables included in the regression analysis between 2002 and 2022 is not included.

Our sample is also restricted by the definition of peer group, which is similar to the one defined in Portugal et al. (2022). It is limited to workers that, in a given year, have a common occupation code and work in the same establishment.[5] Regarding the size of the peer groups,

4. According to the Portuguese Labor Code, Article 275, paragraph 1, clause (a), the monthly wage can be 80% of the national minimum wage if the worker is an apprentice or an intern.

5. In their study, Portugal et al. (2022) use a stricter definition of peer group, using job-title instead of occupation.

since the objective of this work is to estimate the effect of native and immigrant peers' quality on native and immigrant workers' wages, these are restricted to those groups that contain at least two natives and two immigrants, such that each native or immigrant worker has at least one native peer and one immigrant peer. We also restrict peer group size to a maximum of 30 workers to ensure that workers within each group actually engage with each other. Lastly, to guarantee the comparability between fixed affects and to apply the high-dimensional fixed effects model, the final dataset corresponds to a subset of peer groups connected by worker mobility, i.e., the largest connected set, as suggested by ABOWD, CREECY, and KRAMARZ (2002). This last data set covers 83.78% of the observations from our working data set.

Overall, the largest connected set is composed by 1,012,748 observations on 461,694 workers (67.24% natives, and 32.76% immigrants), 76,289 establishments/year, and 3,642 occupations/year. This is reflected in 89,149 peer groups, with an average of 11.36 workers per group (Table A.1 in the Appendix).

III.C Concepts and Wage Setting in Portugal

In the context of this study, as mentioned in the previous subsection, a peer group is defined as all the workers that, in a given year, share the same physical workplace and possess the same occupation. In *QP*, a worker is considered a wage earner identified through a unique identification code, which is the same every year that the worker is part of the observations. The physical workplace is identified through a unique establishment identification code. While a firm refers to the legal entity, a establishment refers to the physical workplace, meaning that each firm code can be associated with more than one establishment code, but each establishment code can only be associated with one firm code. Finally, the occupation of each worker is defined as the professional category of the National Classification of Professions (*Classificação Nacional de Profissões*), where each category corresponds to a unique four-digit code. In terms of immigrant status, it is important to mention that we consider an immigrant as someone that does not have Portuguese nationality. Due to the lack of information regarding a worker's place of birth, there might exist workers identified as immigrants that have already been born in Portugal.

We identify peer effects using wages as our outcome variable, in parallel with other studies. As reported by many authors, the Portuguese labor market is an extreme example of downward wage rigidity (Elsby and Solon 2019; Behr and Pötter 2009; Holden and Wulfsberg 2008). For example, Carneiro, Portugal, and Varejão (2013) showed that, between the years of 1987 and 2009, there were virtually no cuts in nominal wages in Portugal, and Cabral and Duarte (2014) verified that nationality does not substantially influence this wage rigidity. In fact, the Portuguese Labor Code explicitly forbids nominal wage cuts, except for particular cases provided in the law or in collective bargaining agreements.[6] This institutional framework in Portugal can negatively affect the mechanism through which workers' wages internalize the spillover effects from their peers (peer quality). In that respect, we use a more flexible definition of wage, by including all the available components for earnings in *QP*, so that we can capture productivity changes also through these additional components in the total wage, and then divide that by the number of hours worked.[7] Finally, we use real wages to account for adjustments resulting from changes in the inflation rate.

IV Exploratory Analysis

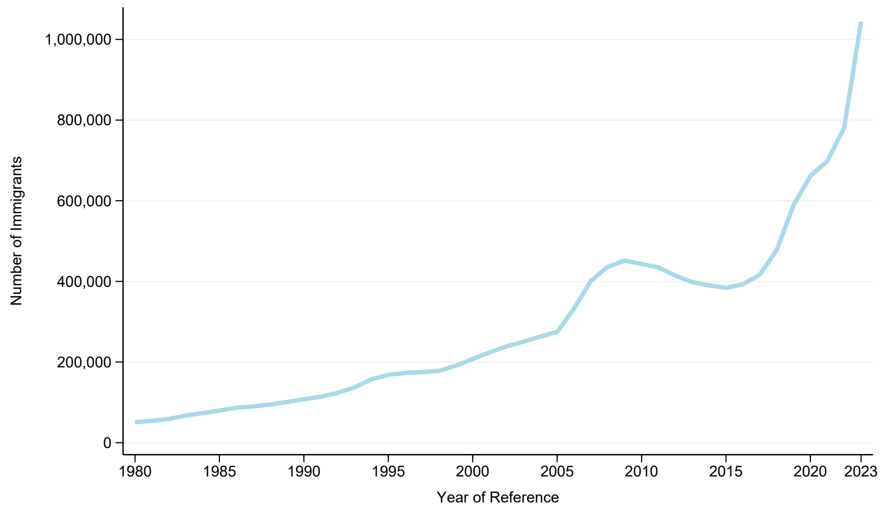
Traditionally, Portugal has been a country of large emigration flows, especially due to work-related reasons and to regions such as Continental Europe and the Americas (particularly to France and Brazil). It was only after the 1974 Carnation Revolution and, consequently, the end

6. See the Portuguese Labor Code, Article 129, paragraph 1, clause (d).
7. As mentioned in Section III.A, this includes the base wage, regular and non-regular benefits, and overtime payments.

of the Portuguese Colonial War, that the traditional emigrant flows from Portugal started to be counterbalanced by very large immigration flows into Portugal.[8]

Since the beginning of the 1980s, the stock of immigrants in Portugal has been steadily increasing (Figure 1), reflecting the process where Portugal is shifting into a destination country for international migrants. For example, in 2023 the number of immigrants in Portugal (around 1,040,000) was 2.36 times higher than in 2010, and 5.03 times higher than in 2000. Although, the composition of these flows is not constant. In the beginning, immigrants arriving into Portugal were coming in large part from Portuguese-speaking regions (mainly countries in Africa, and Brazil). More recently, these flows arrive from different countries such as Ukraine, Moldova, China, India, and Pakistan.

FIGURE 1. FOREIGN POPULATION WITH LEGAL STATUS OF RESIDENCE



Notes: The figure plots the stock of foreign population with legal status of residence in Portugal between 1980 and 2023. It does not include regular foreigners under the granting of stay permits, short-term visas, study, work or temporary stay visas, as well as irregular foreigners.
Source: Statistics Portugal, 1980-2023.

Table A.2 in the Appendix presents descriptive statistics of the largest connected set used in our estimations, covering mean statistics for both natives and immigrants in terms of hourly

8. See, e.g., CARRINGTON and DE LIMA (1996) and BOHNET, PERALTA, and SANTOS (2024) for a detailed analysis of the *retornados* influx into Portugal after the end of the Portuguese Colonial War in 1974 and its consequences to the Portuguese labor market.

total wage, minimum wage earners, gender, age, tenure, schooling, and sector of activity.

Regarding hourly real wage, not controlling for any factor, immigrants in Portugal receive, on average, a lower real wage per hour than natives in the referring period, with the gap being equal to -6.94 log points (times 100). This is consistent with the percentage of workers receiving the minimum wage, accounting for 2.34% of natives and 4.38% of immigrants.

With respect to gender, the percentage of female workers among natives is also higher than among immigrants (41.61% and 38.05%, respectively).

Although presenting just a residual difference, in what concerns the workers' age, immigrants are, on average, younger than natives.

As to tenure, since this tendency of high immigration flows is relatively recent, the number of years of experience of immigrant workers is, on average, substantially lower than that of native workers. While native workers present an average of 6.05 years of experience, immigrant workers have an average of 2.96 years in the period under analysis.

Regarding schooling, the number of years of education is, on average, almost the same for both natives and immigrants.

Finally, in terms of main sectors of activity, immigrants are concentrated in a few number of industries, mainly hotels and restaurants (31.26%), construction (20.97%), and wholesale and retail trade (11.67%). These three sectors account for 63.90% of the sample of immigrants (compared to 55.34% for natives).

V Estimation Strategy

The goal of this thesis is to quantify the impact of peers' ability, distinguishing workers and peers by their immigrant status, on wages, and estimate their contribution, along with labor market sorting, in the immigrant wage gap. In this setting, the estimation strategy of this study is based on the methodological approach applied by Abowd, Kramarz, and Margolis (1999) and ARCIDIACONO et al. (2012) to estimate the effect of labor market sorting and of peers, and on the GELBACH (2016) conditional decomposition method to decompose the change of the immigrant wage gap between these components.

V.A Base-Specification Regression Model

The first step of the estimation strategy of this study is to estimate a basic OLS regression to assess the impact of the worker's nationality in the hourly real wages. This first regression includes a set of observed control variables related to workers' individual characteristics, and year fixed effects to control for variables that are constant across all workers but vary over time (e.g., inflation). In this setting, the base-specification regression has the following form:

$$\omega_{it} = X_{it}' \, \beta^{base} + \gamma^{base} \, Immig_i + \tau_t^{base} + \varepsilon_{it} \tag{1}$$

where ω_{it} is the logarithm of the real wage per hour for worker i ($i \in \{1, ..., N\}$) at year t ($t \in \{1, ..., T\}$); $Immig_i$ is a dummy variable equal to 1 when worker i is an immigrant or 0 when worker i is a native; X_{it}' is a set containing other observed workers' individual characteristics;[9] τ_t is the year fixed effect; and ε_{it} is the stochastic error term that is assumed to follow the conventional assumptions, such as strict exogeneity. The results from regression (1) are presented in Table 1.

V.B Full-Specification Regression Model

The estimation of peer effects raises a number of challenges that are already well known in the literature. These problems are (i) the reflection problem, (ii) sorting and homophily, and (iii) the existence of unobserved correlated shocks.

First, the reflection problem, introduced by MANSKI (1993), arises when one tries to assess

9. In more detail, this set includes the variables age, age squared, tenure, tenure squared, gender and schooling.

if the average behavior of a group is influencing the behavior of its individual members. For example, imagine the following specification: $y_{ig} = \alpha + \beta\, \overline{y}_g + \xi_{ig}$, where i corresponds to the worker in peer group g; y_{ig} corresponds, for example, to the wage of worker i in peer group g; $\overline{y}_g$ corresponds to the average wage of peer group g; and ξ_{ig} is the respective error term. The objective of this regression would be to estimate the peer effect through coefficient β. However, regressing an individual outcome on a group mean outcome will always lead to a coefficient of one.[10] To circumvent the reflection problem, we adopt the framework developed by Arcidiacono et al. (2012) to estimate peer effects. The authors' strategy consists in including the average fixed effect of peers computed through a linear combination of individual fixed effects as a covariate in the regression, whose coefficient reflects the impact of peers' quality within the group.

Second, the concern of sorting corresponds to the non-random allocation of workers across establishments, occupations, or peer groups, while homophily is the tendency of workers within these groups to display similar behaviors and outcomes due to shared characteristics and environments (and not because of their social interactions). In order to prevent these two obstacles, our study extends the canonical AKM model, incorporating a set of fixed effects. To account for the possible non-random sorting of high-ability workers into high-ability peer groups we include worker fixed effects, which also avoids the problem of homophily. Additionally, we introduce establishment/occupation fixed effects, a multi-dimensional effect that allows us to control for the potential sorting of high-ability workers into establishments, occupations, or combinations establishment/occupation offering higher wages.

Third, the issue of unobserved correlated shocks is associated with the possible correlation between changes in unobserved background characteristics at the peer group level (i.e., at the establishment/occupation level) and changes in the quality of peers. If these unobserved back-

10. See Angrist (2014) for the discussion of the OLS properties of regression used in this example.

ground characteristics alter at the peer group/time level, and consequently change peer quality between two consecutive periods, the estimator of peer effects will be biased. In HONG and LATTANZIO (2022), a case is presented where a firm decides to implement an automation process that complements white-collar workers but substitutes blue-collar workers. If we assume that white-collar workers are more skilled, this investment would lead to an increase of peer quality and productivity (and, consequently, wages) at the same time, leading to a positive bias of the peer effect estimator. Having this in mind, so as to account for time-variant unobserved background characteristics, we expand the previous establishment/occupation fixed effect to include the time (year) effect. Nevertheless, by including establishment/occupation/year fixed effects, the identification of peer effects can only be done through changes in peer group sizes.[11] In the end, although we have now only one way to identify peer effects, done trough changes in the peer group size, we are able to take into account the effect of changes in wage frameworks in each firm and across different establishments of the same firm over time, the effect in each occupation over time, and the interaction between effects of establishment, occupation and year.

As mentioned before, this study incorporates the methodology of ARCIDIACONO et al. (2012). We then extend it by allowing for both heterogeneity in the response to peers and heterogeneity in the influence of peers. In other words, we allow for the peer effect to be differentiated into four types of worker/co-worker relations in the workplace concerning nationality: (i) the effect of native peers on native workers; (ii) the effect of immigrant peers on native workers; (iii) the effect of native peers on immigrant workers; and (iv) the effect of immigrant peers

11. See the discussion, e.g., in CORNELISSEN, DUSTMANN, and SCHÖNBERG (2017) and PORTUGAL et al. (2022). More precisely, the average fixed effect of a certain peer group can de described as the difference between two components: the ratio between the sum of the fixed effect of peers in the group and the number of peers, and the ratio between the worker fixed effect and the number of peers in the group. If the number of peers in the group stays constant over time, the first component will be fully captured by the establishment/occupation/year fixed effect (i.e., peer group level fixed effect), while the second will be fully captured by the worker fixed effect.

on immigrant workers. Then, our full-specification model can be described as:

$$\omega_{it} = X'_{it}\,\beta^{full} + \gamma^{full}\,Immig_i + \alpha_i + \delta_1\,\overline{\alpha}^{NN}_{-it} + \delta_2\,\overline{\alpha}^{NI}_{-it} + \delta_3\,\overline{\alpha}^{IN}_{-it} + \delta_4\,\overline{\alpha}^{II}_{-it} + \pi_{eot} + \epsilon_{it} \quad (2)$$

where α_i is the worker fixed effect; $\overline{\alpha}^{NN}_{-it}$ is the average of the fixed effects for the native co-workers of the native worker i at time t; $\overline{\alpha}^{NI}_{-it}$ is the average of the fixed effects for the immigrant co-workers of native worker i at time t; $\overline{\alpha}^{IN}_{-it}$ is the average of the fixed effects for the native co-workers of the immigrant worker i at time t; $\overline{\alpha}^{II}_{-it}$ is the average of the fixed effects for the immigrant co-workers of the immigrant worker i at time t; π_{eot} is the establishment/occupation/year fixed effect; and ϵ_{it} is the error term. In more detail, the average fixed effects just mentioned are estimated as it follows:

$$\overline{\alpha}^{NN}_{-it} = \frac{\sum_{j\in\mathbb{M}^N_{gt}}(\alpha_j) - \alpha_i}{|\mathbb{M}^N_{gt}|-1}\ , i \in \mathbb{M}^N_{gt} \quad ; \quad \overline{\alpha}^{NI}_{-it} = \frac{\sum_{j\in\mathbb{M}^I_{gt}}\alpha_j}{|\mathbb{M}^I_{gt}|}\ , i \notin \mathbb{M}^I_{gt}$$

$$\overline{\alpha}^{IN}_{-it} = \frac{\sum_{j\in\mathbb{M}^N_{gt}}\alpha_j}{|\mathbb{M}^N_{gt}|}\ , i \notin \mathbb{M}^N_{gt} \quad ; \quad \overline{\alpha}^{II}_{-it} = \frac{\sum_{j\in\mathbb{M}^I_{gt}}(\alpha_j) - \alpha_i}{|\mathbb{M}^I_{gt}|-1}\ , i \in \mathbb{M}^I_{gt}$$

$$\text{for } i \in \mathbb{M}_{gt}\ ; \text{ and } \mathbb{M}^N_{gt}, \mathbb{M}^I_{gt} \subset \mathbb{M}_{gt}$$

where $\mathbb{M}_{gt}$ is the set of individuals in peer group g at time t; $\mathbb{M}^N_{gt}$ is the subset of native individuals in peer group g at time t; $\mathbb{M}^I_{gt}$ is the subset of immigrant individuals in peer group g at time t; $\sum_{j\in\mathbb{M}^N_{gt}}(\alpha_j)-\alpha_i$ and $\sum_{j\in\mathbb{M}^I_{gt}}(\alpha_j)-\alpha_i$ correspond to the sum of the individual fixed effects of all workers j belonging to the subgroup of natives and immigrants, respectively, of peer group g at time t minus the individual fixed effect of worker i belonging to that same subgroup.[12] The results from regression (2) are presented in Table 2, Column (2).

In order to estimate this model, we use an iterative approach that consists in alternating between the estimation of the individual fixed effects and the coefficient associated with the peer effect, such that these estimates converge to an imposed criterion of precision.[13] It is also

12. $|-|$ corresponds to the number of individuals in the corresponding subset.
13. This iterative process is applied using the *regpeer* ado-file developed by Portugal et al. (2022).

important to note that, when adding the individual worker fixed effects, the OLS estimation does not imply that time-invariant variables related to the worker (such as nationality or gender) are fully absorbed by the fixed effect, since it is no longer necessary for the fixed effect to be orthogonal to the error term.[14] Although, this makes the coefficient of these variables more difficult to interpret, since we are not able to separate the individual effect from the average effect.

V.C Decomposition of the Immigrant Wage Gap

In the last step of our estimation strategy we apply the conditional decomposition method developed by GELBACH (2016). The method uses the OLS basic formulation for the issue of omitted variables bias to assess how much of the change in the coefficient of interest from a base- to a full-specification model is attributed to a set of additional controls added when moving from the restricted to the unrestricted specification.

In our particular case, we use the Gelbach's decomposition method to identify how much of the difference between the coefficient of nationality in the base-specification model (1) and the coefficient of nationality in the full-specification model (2) is attributed to each of the additional fixed effect added in the full model, i.e., the worker fixed effects (α_i), the establishment/occupation/year fixed effects (π_{eot}), and each variable of the human capital spillovers ($\overline{\alpha}^{NN}_{-it}$, $\overline{\alpha}^{NI}_{-it}$, $\overline{\alpha}^{IN}_{-it}$, $\overline{\alpha}^{II}_{-it}$).

This method relies on a set of auxiliary regressions that regress each of the decomposition components on the control variables used in the base-specification model. In the case of this study, each auxiliary regression uses the estimated fixed effects that were just described and that were estimated in the full-regression model as the dependent variable, and the variables used in the base-specification model (nationality, other workers' individual characteristics, and year

14. See the discussion in PORTUGAL et al. (2022).

fixed effects) as explanatory variables. The general specification of this auxiliary regression can be written as follows:

$$\hat{\mathbb{Y}} = X'_{it}\,\beta^{\mathbb{Y}} + \phi^{\mathbb{Y}}\,Immig_i + \tau_t^{\mathbb{Y}} + \upsilon_{it}^{\mathbb{Y}} \tag{3}$$

where $\hat{\mathbb{Y}} \in \{\hat{\alpha}_i, \hat{\overline{\alpha}}_{-it}^{NN}, \hat{\overline{\alpha}}_{-it}^{NI}, \hat{\overline{\alpha}}_{-it}^{IN}, \hat{\overline{\alpha}}_{-it}^{II}, \hat{\pi}_{eot}\}$.

Then, it is possible to assess how each of the fixed effects affect the difference between $\hat{\gamma}^{base}$ and $\hat{\gamma}^{full}$ by using the estimated coefficients $\hat{\phi}$ from each auxiliary regression (3) and the fixed effects estimations from the full-specification model (2):

$$\hat{\gamma}^{base} - \hat{\gamma}^{full} = \hat{\phi}^{\,\alpha} + \hat{\phi}^{\,\pi} + \hat{\delta}_1\hat{\phi}^{\,\overline{\alpha}^{NN}} + \hat{\delta}_2\hat{\phi}^{\,\overline{\alpha}^{NI}} + \hat{\delta}_3\hat{\phi}^{\,\overline{\alpha}^{IN}} + \hat{\delta}_4\hat{\phi}^{\,\overline{\alpha}^{II}} \tag{4}$$

Each component describes how much of the change in the immigrant wage gap is explained by the respective fixed effect, which does not depend on the order in which covariates are added to the regression. The results from the Gelbach's conditional decomposition are presented in Table 3, Columns (3)-(8).

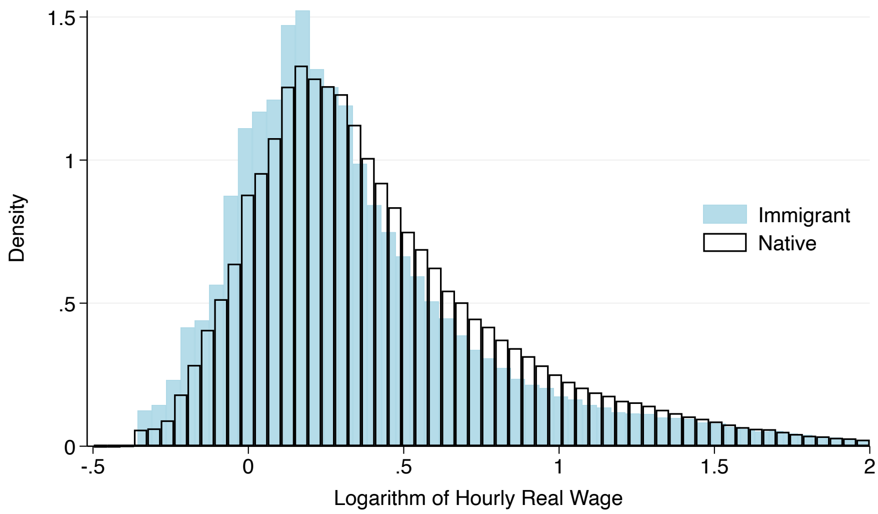
The establishment/occupation/year fixed effect can be further decomposed to estimate the individual contributions of establishments and occupations over time. Following the approach of Portugal et al. (2022), we assume that establishment and occupation effects are orthogonal to their interactions with time. Accordingly, we perform a linear regression of the fitted values of the establishment/occupation/year fixed effect on establishment/year and occupation/year fixed effects, controlling for the same variables included in the base specification model. This approach yields estimates for the individual contributions of establishments and occupations over time, with the residual component capturing the interactions between establishments, occupations, and time. The results are presented in Table 3, Columns (10)-(12).

VI Empirical Results

VI.A Base-Specification Analysis

As noted in Section IV, the difference in the logarithm of hourly real wages between immigrants and natives, without controlling for any other characteristic, is -6.94 log points. By observing its distribution for natives and immigrants (Figure 2), it is noticeable that the distribution for immigrant workers is slightly to the left of the natives' distribution.

FIGURE 2. HISTOGRAM OF THE DISTRIBUTION OF HOURLY REAL WAGES FOR NATIVES AND IMMIGRANTS



Notes: The figure plots the histogram of the logarithm of hourly real wages, trimmed between the values of -0.5 and 2.
Source: Quadros de Pessoal, 2002-2022.

In order to disentangle the immigrant wage gap into the different components, we first have to define a base regression, where this gap is controlled for several observable characteristics of the worker, as detailed in Section V.A. As such, we start by estimating the regression specified in (1), reporting the results for all parameters in Table 1.

As expected, the coefficient associated with the immigrant status, when controlling for other individual characteristics, is negative, i.e., on average, an immigrant worker receives 5.53% (5.70 log points) less than a native worker. Also, in this estimation, wages increase at a decreasing rate for both age and tenure. For age and tenure, the maximum is reached at around 49 and 46 years, respectively. In Portugal, the gender wage gap is estimated to be 25.06% (22.36 log points) during this period. Finally, each additional year of education increases wages by around 4.57% (4.46 log points).

TABLE 1. BASE-SPECIFICATION MODEL ESTIMATES

	Base Specification
Age	0.0264
	(0.0004)
Age Squared	-0.0003
	(0.0000)
Tenure	0.0161
	(0.0002)
Tenure Squared	-0.0002
	(0.0000)
Gender (Male=1)	0.2236
	(0.0015)
Schooling	0.0446
	(0.0003)
Nationality (Immigrant=1)	-0.0569
	(0.0016)
Year Effects (τ_t)	✓
Observations	1,012,748
R-Squared	0.1925

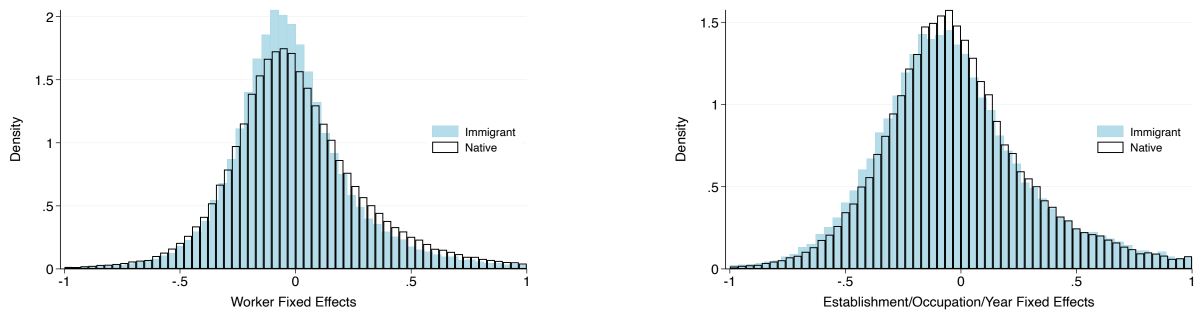
Notes: The table reports the estimates for the base-specifications model in regression (1). The dependent variable is the logarithm of hourly real wages. The variables of age, tenure and schooling are measured in years. Standard errors are in parentheses and clustered at the worker level.
Source: Quadros de Pessoal, 2002-2022.

VI.B Full-Specification Analysis

We now present the estimates of the full-specification model that, as explained in Section V.B, includes establishment/occupation/year fixed effects and the average fixed effects of immigrant/native peers for immigrant/native workers, in line with the iterative procedure introduced by ARCIDIACONO et al. (2012). The results from regression (2) are reported in Table 2, Column (2).

The distribution of worker and establishment/occupation/year fixed effects highlights the differences between immigrants and natives (Figure 3). The distribution for natives is broader, with a higher density at higher values compared to immigrants. A similar pattern is evident in the distribution of peer group fixed effects.

FIGURE 3. HISTOGRAM OF WORKER AND PEER GROUP FIXED EFFECTS



(A) WORKER

(B) PEER GROUP

Notes: The figures plot the histograms of worker and peer group (establishment/occupation/year) fixed effects. The figure is trimmed between the values of -1 and 1.
Source: *Quadros de Pessoal*, 2002-2022.

As previously stated, the interpretation of coefficients of time-invariant variables in the model developed by ARCIDIACONO et al. (2012) is more complex, since it is not possible to separate the individual and the average effects. One possible interpretation of the coefficient associated to the immigrant status is that, in the absence of spillovers, on average, the wage of an immigrant is 0.22% lower than that of a native, taking already into consideration worker heterogeneity and labor market sorting.

Regarding time-variant variables, as in the base-specification model, wages increase at a decreasing rate for age and tenure, with their maximums being reach at 53 and 19 years, respectively.

Concerning the effects of human capital spillovers, we report the estimates for the effect of average peer quality, computed as the average worker fixed effect of coworkers in the peer group. One of the core contributions of this study is the quantification of this effect while allowing for heterogeneous responses (native vs. immigrant workers) and heterogeneous influences (native vs. immigrant coworkers). The distribution of the estimated values for the human capital spillovers are illustrated in Figure A.1 in the Appendix.

We obtain that, when both worker and peer are natives, a 10% increase in the average peer quality increases wages by 0.94%, on average, ceteris paribus. In other words, a one standard deviation increase in the measure of human capital spillovers of native peers on native workers (0.2341), leads to wage increase of around 2.21%.[15] On the opposite, when both worker and peer are immigrants, we have that a 10% increase in the average peer quality increases wages by 0.79%, or, to put it differently, a one standard deviation increase in the human capital of immigrant peers on immigrant workers (0.1493) results in an increase on the wage of these workers of around 1.18%.

When looking to heterogeneous relations in the workplace, i.e., native (immigrant) workers and immigrant (native) peers within each group, the results differ. In the case where there is a native worker and an immigrant peer in the same group, a 10% increase of the immigrant peers' quality leads to a decrease in natives' wages by 0.49%. In the same way, a one standard deviation increase (0.2293) results in a decrease of natives' wages by 1.12%. On the contrary, when there is an immigrant worker and native peers, a 10% increase of the human capital of native peers leads to an increase in immigrants' wages by 0.41%, or an increase of one standard deviation (0.1490) in this measure leads to an increase of immigrants' wages of about 0.61%. Nevertheless, it is important to notice that these last two effects are not statistically significant.

It can also be observed that when both the worker and their peers belong to the same nationality-specific group (either native or immigrant), the impact of peers' average ability is greater than when the worker and peers differ in their immigrant status ($\hat{\delta}_1 > \hat{\delta}_4 > \hat{\delta}_3 > \hat{\delta}_2$).

15. Computed as (0.0943 x 0.2341). Table A.3 in the Appendix reports this and other relevant statistics of peer quality, i.e., the fixed effects of each peer.

TABLE 2. BASE- AND FULL-SPECIFICATION MODELS ESTIMATES

	Base Specification	Full Specification
	(1)	(2)
Age	0.0264	0.0225
	(0.0004)	(0.0013)
Age Squared	-0.0003	-0.0002
	(0.0000)	(0.0000)
Tenure	0.0161	0.0092
	(0.0002)	(0.0004)
Tenure Squared	-0.0002	-0.0002
	(0.0000)	(0.0000)
Gender (Male=1)	0.2236	0.0331
	(0.0015)	(0.0324)
Schooling	0.0446	0.0009
	(0.0003)	(0.0004)
Nationality (Immigrant=1)	-0.0569	-0.0022
	(0.0016)	(0.0026)
Effect on Native Worker:		
- HC Spillovers of Native Co-Worker ($\overline{\alpha}_{-it}^{NN}$)	-	0.0901
		(0.0234)
- HC Spillovers of Immigrant Co-Worker ($\overline{\alpha}_{-it}^{NI}$)	-	-0.0501
		(0.0337)
Effect on Immigrant Worker:		
- HC Spillovers of Native Co-Worker ($\overline{\alpha}_{-it}^{IN}$)	-	0.0404
		(0.0329)
- HC Spillovers of Immigrant Co-Worker ($\overline{\alpha}_{-it}^{II}$)	-	0.0763
		(0.0212)
Year Effects (τ_t)	✓	
Worker Effects (α_i)		✓
Establishment/Occupation/Year Effects (π_{eot})		✓
Observations	1,012,748	1,012,748
R-Squared	0.1925	0.9440

Notes: The table reports the estimates for the base- and full-specifications models. The dependent variable in each model is the logarithm of hourly real wages. The variables of age, tenure and schooling are measured in years. Column (1) presents the regression estimates obtained from the base-specification model (1). Column (2) presents the regression estimates obtained from the full-specification model (2). Standard errors are in parentheses and, in Column (1), they are clustered at the worker level, while in Column (2) they are clustered at the peer group level.
Source: Quadros de Pessoal, 2002-2022.

This may highlight the significance of shared cultural backgrounds and experiences. Shared cultural norms, values, and practices likely foster better communication and mutual understanding, which in turn enhance collaboration and knowledge sharing within teams. Workers with similar cultural backgrounds may also find it easier to coordinate efforts and develop stronger social bonds, all of which may help for effective teamwork and productivity. Moreover, cultural familiarity could reduce friction that might arise from differences in work styles or behavioral norms, enabling smoother interactions and a more cohesive work environment. However, it is important to note that the immigrant group encompasses a wide range of nationalities, adding considerable heterogeneity that could influence the magnitude of the effect.

VI.C Decomposition of the Immigrant Wage Gap

Finally, in order to understand the contribution of the worker, establishment/occupation/year, and peer quality components to the wage differential between immigrants and natives, we use the decomposition method developed by GELBACH (2016). As discussed earlier, this method allows to quantify the share of the variation attributed to each of components mentioned.

Table 3 presents the results of this decomposition. It indicates that the difference between the immigrant wage gap in the base-specification model (-0.0569) and in the full-specification model (-0.0022) can be decomposed in six components: the portion associated to the sorting across different establishment/occupation combinations over time (-0.0286), the contribution of the worker component (-0.0237), the human capital spillovers to native workers from native (-0.0010) and immigrant (-0.0004) co-workers, and the human capital spillovers to immigrant workers from native (0.0001) and immigrant (-0.0012) co-workers.

The component of establishment/occupation/year represents more than half of the wage differential between immigrants and natives (52.30%). It reflects that immigrant workers are sorted into establishments and occupations with lower wages. By further decomposing this

TABLE 3. CONDITIONAL DECOMPOSITION OF THE IMMIGRANT WAGE GAP

Panel A - Gelbach Decomposition of the Immigrant Wage Gap

		Decomposition into:					
				HC Spillovers to Native Worker of:		HC Spillovers to Immigrant Worker of:	
Base Specification	Full Specification	Establishment/Occupation/Year	Worker	Native Co-Worker	Immigrant Co-Worker	Native Co-Worker	Immigrant Co-Worker
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
-0.0569	-0.0022	-0.0286	-0.0237	-0.0010	-0.0004	0.0001	-0.0012
(0.0016)	(0.0002)	(0.0014)	(0.0013)	(0.0001)	(0.0001)	(0.0000)	(0.0001)

Panel B - Decomposition of the Establishment/Occupation/Year Fixed Effect

	Decomposition into:		
Establishment/Occupation/Year	Establishment/Year	Occupation/Year	Interaction
(9)	(10)	(11)	(12)
-0.0286	-0.0260	-0.0019	-0.0007
(0.0014)	(0.0015)	(0.0014)	(0.0001)

Notes: The table reports the estimates for the conditional decomposition of the immigrant coefficient, based on the method developed by GELBACH (2016). Panel A refers to the decomposition of the immigrant wage gap, and Panel B refers to the decomposition of the establishment/occupation/year fixed effect. The nationality coefficients of the base- and full-specification models in Column (1) and (2) are derived from regressions (1) and (2), respectively. The conditional decomposition results in Columns (3)-(8) are derived from equation (4). Standard errors are in parentheses and clustered at the group level of the corresponding fixed effect.
Source: Quadros de Pessoal, 2002-2022.

effect, we can observe that immigrants are sorted into worse paying establishments (-0.0260), and into worse paying occupations (-0.0019), with the first dominating the aggregated fixed effect. The remaining contribution (-0.0007) is residual and attributed to the interaction between the two effects. Similarly, the worker component also expresses a large portion of the immigrant wage gap, around 43.23%. It can be interpreted as the wage gap of immigrants if workers are randomly assigned into establishment, occupations, and peer groups.

Finally, the contribution of the different human capital spillovers specifications to the immigrant wage gap reveals that their combined impact is relatively minor, accounting for only 4.47% of the total gap. The individual contributions of these spillovers illustrate how the wage gap would change if such effects were absent. For instance, eliminating the human capital spillovers from native co-workers to native workers would reduce the immigrant wage gap by 0.1 percentage points. From another perspective, this specific spillover channel accounts for 1.88% of the immigrant wage gap. This could be attributed to the fact that the positive impact of high-quality native peers on native workers amplifies the wage disparity by increasing native workers' earnings, thereby widening the gap between natives and immigrants. Similar findings are observed for human capital spillovers from immigrant co-workers to immigrant workers (-0.0012), which contribute 2.12% to the immigrant wage gap. However, removing this type of spillover would widen the wage gap because higher-quality immigrant peers positively influence immigrants' wages, and without this effect, immigrant wages would decrease, thereby increasing the wage disparity between immigrants and natives. The remaining two specifications of human capital spillovers contribute only marginally to the immigrant wage gap.

VI.D Alternative Specifications

With the intention of testing different specifications of our full-estimation model, we undertake three different exercises based on alternative levels of controls for heterogeneity, through different fixed effect specifications. The results from this exercise are presented in Table A.4 in the Appendix, Columns (2)-(4). We observe that more aggregated fixed effects specifications yield higher peer effects coefficients. In our view, this can largely be attributed to the looser control over sorting phenomena in these alternative specifications, and to different requirements of parameter identification.

VII CONCLUSION

In this study, we analyzed the impact of immigrant and native peers' ability on immigrant and native workers' wages. We used a matched employer-employee data set, and applied an estimation strategy that is able to estimate peer effects while circumventing problems of endogeneity.

We found that when both worker and peer are natives or immigrants, a one standard deviation increase in the average quality of peers increases wages by 2.21% or 1.18%, respectively. On the other hand, when a one standard deviation increase in the average quality of immigrant peers leads to a decrease of natives' wages by 1.12%, while a one standard deviation increase in the average quality of native peers leads to an increase of immigrants' wages by 0.61%. These last two effects, however, are not statistically significant.

Finally, we examine how these peer effects, combined with worker, establishment, occupation, and year components, contribute to the immigrant wage gap. We found that the establishment/occupation/year component explains 52.30% of the wage differential between immigrants and natives, meaning that immigrants are sorted into establishments and occupations offering lower wages. Then, it is followed by the contribution of the worker component (43.23%), and peer effects (4.47%).

Our results may reflect the importance of shared cultural backgrounds and experiences, where common values and practices likely foster better communication and mutual understating, contributing to a better knowledge sharing and increases in productivity.

References

ABOWD, JOHN M., ROBERT H. CREECY, and FRANCIS KRAMARZ. 2002. *Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data.* Longitudinal Employer-Household Dynamics Technical Papers 2002-06. Center for Economic Studies, U.S. Census Bureau.

ABOWD, JOHN M., FRANCIS KRAMARZ, and DAVID N. MARGOLIS. 1999. "High Wage Workers and High Wage Firms." *Econometrica* 67 (2): 251–333.

ANGRIST, JOSHUA D. 2014. "The Perils of Peer Effects." *Labour Economics* 30:98–108.

ARCIDIACONO, PETER, GIGI FOSTER, NATALIE GOODPASTER, and JOSH KINSLER. 2012. "Estimating Spillovers Using Panel Data, With an Application to the Classroom." *Quantitative Economics* 3 (3): 421–470.

ARROW, KENNETH. 1971. *The Theory of Discrimination.* Working Papers 403. Princeton University, Department of Economics, Industrial Relations Section.

AYDEMIR, ABDURRAHMAN, and MIKAL SKUTERUD. 2005. "Explaining the Deteriorating Entry Earnings of Canada's Immigrant Cohorts, 1966 – 2000." *Canadian Journal of Economics/Revue canadienne d économique* 38 (2): 641–672.

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

BECKER, GARY S. 1957. *The Economics of Discrimination.* Chicago: The University of Chicago Press.

BEHR, ANDREAS, and ULRICH PÖTTER. 2009. "Downward Wage Rigidity in Europe: A New Flexible Parametric Approach and Empirical Results." *German Economic Review* 11 (2): 169–187.

BOHNET, LARA, SUSANA PERALTA, and JOÃO PEREIRA DOS SANTOS. 2024. "Cousins From Overseas: How the Existing Workforce Adapts to a Massive Forced Return Migration Shock." *European Economic Review,* 104925.

CABRAL, SÓNIA, and CLÁUDIA DUARTE. 2013. "Mind the gap! The relative wages of immigrants in the Portuguese labour market." *Working Papers 05 - 2013, Banco de Portugal.*

———. 2014. "Nominal and Real Wage Rigidity: Does Nationality Matter?" *IZA Journal of European Labor Studies* 3 (1).

CARNEIRO, ANABELA, NATÉRCIA FORTUNA, and JOSÉ VAREJÃO. 2012. "Immigrants at New Destinations: How They Fare and Why." *Journal of Population Economics* 25 (3): 1165–1185.

Carneiro, Anabela, Pedro Portugal, and José Varejão. 2013. "Catastrophic Job Destruction During the Portuguese Economic Crisis." *Journal of Macroeconomics* 39:444–457.

Carrington, William J., and Pedro J. F. De Lima. 1996. "The Impact of 1970s Repatriates From Africa on the Portuguese Labor Market." *ILR Review* 49 (2): 330–347.

Chan, Tat Y., Jia Li, and Lamar Pierce. 2014. "Compensation and Peer Effects in Competing Sales Teams." *Management Science* 60 (8): 1965–1984.

Chiswick, Barry R. 1978. "The Effect of Americanization on the Earnings of Foreign-born Men." *Journal of Political Economy* 86 (5): 897–921.

Chiswick, Barry R., and Paul W. Miller. 2015. *International Migration and the Economics of Language,* 211–269.

Cohen-Goldner, Sarit, and Zvi Eckstein. 2002. "Labor Mobility of Immigrants: Training, Experience, Language and Opportunities." *IZA Discussion Papers* 159.

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

De Matos, Ana Damas. 2016. "Firm Heterogeneity and Immigrant Wage Assimilation." *Applied Economics Letters* 24 (9): 653–657.

Elsby, Michael W. L., and Gary Solon. 2019. "How Prevalent Is Downward Rigidity in Nominal Wages? International Evidence From Payroll Records and Pay Slips." *The Journal of Economic Perspectives* 33 (3): 185–201.

Falk, Armin, and Andrea Ichino. 2006. "Clean Evidence on Peer Effects." *Journal of Labor Economics* 24 (1): 39–57.

Friedberg, Rachel M. 2000. "You Can't Take It With You? Immigrant Assimilation and the Portability of Human Capital." *Journal of Labor Economics* 18 (2): 221–251.

Gelbach, Jonah B. 2016. "When Do Covariates Matter? And Which Ones, and How Much?" *Journal of Labor Economics* 34 (2): 509–543.

Guryan, Jonathan, Kory Kroft, and Matthew J Notowidigdo. 2009. "Peer Effects in the Workplace: Evidence From Random Groupings in Professional Golf Tournaments." *American Economic Journal Applied Economics* 1 (4): 34–68.

Holden, Steinar, and Fredrik Wulfsberg. 2008. "Downward Nominal Wage Rigidity in the OECD." *The B E Journal of Macroeconomics* 8 (1).

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

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

KAUR, SUPREET, MICHAEL KREMER, and SENDHIL MULLAINATHAN. 2010. “Self-Control and the Development of Work Arrangements.” *American Economic Review* 100 (2): 624–628.

LINDQUIST, MATTHEW J., JAN SAUERMANN, and YVES ZENOU. 2015. “Network Effects on Worker Productivity.” *CEPR Discussion Papers.*

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

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

MARSHALL, ALFRED. 1890. *Principles of Economics.*

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

PHELPS, EDMUND S. 1972. “The Statistical Theory of Racism and Sexism.” *The American Economic Review* 62 (4): 659–661.

PORTUGAL, PEDRO, HUGO REIS, PAULO GUIMARÃES, and ANA RUTE CARDOSO. 2022. “Human Capital Spillovers and Returns to Education.” *Working Papers 19 - 2022, Banco de Portugal.*

APPENDIX

TABLE A.1. SAMPLE STATISTICS

Number of observations after sample selection	1,208,816
- Number of connected observations	1,012,748
- Number of unconnected observations	196,068
- Ratio of connected observations	83.78%
Number of workers	461,694
- Share of native workers	67.24%
- Share of immigrant workers	32.76%
Number of establishments/year	76,289
Number of occupations/year	3,642
Number of peer groups	89,149
Average size of peer groups	11.36

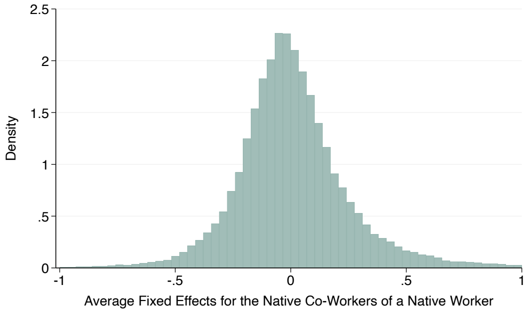
Notes: The table reports sample statistics for the data set after sample selection, as explained in Section III.B, and the largest connected set. The number of connected observations corresponds to a subset of peer groups connected by worker mobility, i.e., the largest connected set. The ratio of connected observations corresponds to the coverage of the largest connected set compared to the number of observations after sample selection. The number of workers, peer groups, establishments/year, and occupations/year refer to the largest connected set.
Source: Quadros de Pessoal, 2002-2022.

TABLE A.2. SUMMARY STATISTICS - AVERAGE 2002-2022

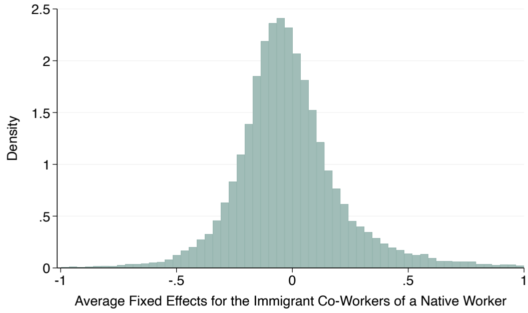
	Mean	Standard Deviation	Minimum	Maximum
Panel A - Native Workers				
Hourly Real Wage	0.4338	0.4524	-0.5354	6.5780
Minimum Wage Earners	0.0234	0.1511	0	1
Gender (Male=1)	0.5839	0.4929	0	1
Age	39.5094	11.6950	18	64
Tenure	6.0510	7.4159	0	49.5
Schooling	8.3722	3.4641	0	23
Main Sectors of Activity				
Hotels and Restaurants Workers	0.2441	0.4295	0	1
Construction Workers	0.1594	0.3660	0	1
Wholesale and Retail Trade Workers	0.1499	0.3570	0	1
Panel B - Immigrant Workers				
Hourly Real Wage	0.3644	0.4562	-0.5279	5.8594
Minimum Wage Earners	0.0438	0.2047	0	1
Gender (Male=1)	0.6195	0.4855	0	1
Age	36.4449	10.0650	18	64
Tenure	2.9558	4.0947	0	47.5
Schooling	8.9390	3.6431	0	23
Main Sectors of Activity				
Hotels and Restaurants Workers	0.3126	0.4635	0	1
Construction Workers	0.2097	0.4071	0	1
Wholesale and Retail Trade Workers	0.1167	0.3211	0	1

Notes: The table reports descriptive statistics for the hourly real wage (in logarithms), minimum wage earners, gender, age, tenure, schooling, and main sectors of activity. A worker is considered to be earning the minimum wage if the nominal total wage lies in an interval of -1/+1 centered in the minimum wage of that year. Panel A presents the statistics for native workers, while Panel B presents them for immigrant workers. All statistics refer to the largest connected set.
Source: Quadros de Pessoal, 2002-2022.

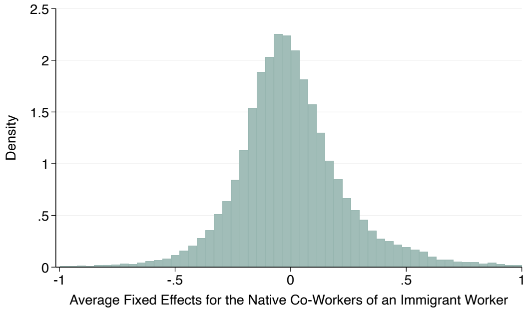
FIGURE A.1. HISTOGRAM OF AVERAGE PEER FIXED EFFECTS



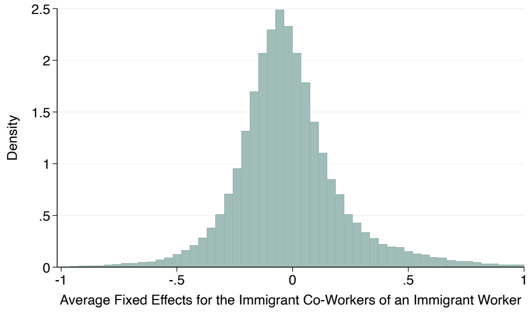
(A) NATIVE WORKER AND NATIVE CO-WORKER



(B) NATIVE WORKER AND IMMIGRANT CO-WORKER



(C) IMMIGRANT WORKER AND NATIVE CO-WORKER



(D) IMMIGRANT WORKER AND IMMIGRANT CO-WORKER

Notes: The figures plot the histograms of the average peer fixed effect for the different workplace interactions between workers and peers in terms of nationality. The figure is trimmed between the values of -1 and 1.
Source: Quadros de Pessoal, 2002-2022.

TABLE A.3. SUMMARY STATISTICS - WORKER, PEER GROUP, AND AVERAGE PEER FIXED EFFECTS

	Mean	Standard Deviation	Minimum	Maximum
Panel A - Native Workers				
Worker Fixed Effects	0.0079	0.3446	-4.5707	4.8780
Establishment/Occupation/Year Fixed Effects	0.0058	0.3905	-3.1400	4.8525
Panel B - Immigrant Workers				
Worker Fixed Effects	-0.0156	0.2990	-4.3161	3.7851
Establishment/Occupation/Year Fixed Effects	-0.0128	0.3903	-3.1400	4.8525
Panel C - Average Peer Fixed Effects				
Effect on Native Worker:				
- HC Spillovers of Native Co-Worker ($\overline{\alpha}_{-it}^{NN}$)	0.0054	0.2341	-4.1511	2.7292
- HC Spillovers of Immigrant Co-Worker ($\overline{\alpha}_{-it}^{NI}$)	-0.0071	0.2293	-3.9930	3.0746
Effect on Immigrant Worker:				
- HC Spillovers of Native Co-Worker ($\overline{\alpha}_{-it}^{IN}$)	0.0007	0.1490	-4.1227	2.6516
- HC Spillovers of Immigrant Co-Worker ($\overline{\alpha}_{-it}^{II}$)	-0.0049	0.1493	-4.3161	3.7851

Notes: The table reports descriptive statistics for the worker fixed effects, peer group (establishment/occupation/year) fixed effects, and average peer fixed effects. Panel A presents the statistics for native workers, Panel B presents them for immigrant workers, and Panel C presents the statistics for the average peer fixed effects considering different workplace interactions between workers and peers in terms of nationality. These statistics are illustrated in Figures 3 and A.1. All statistics were refer to the fixed effect estimates obtained from the full-specification model presented in Section V.B.
Source: Quadros de Pessoal, 2002-2022.

TABLE A.4. PEER EFFECTS ON ALTERNATIVE LEVELS OF HETEROGENEITY

	Baseline	Alternative FE Specifications		
	(1)	(2)	(3)	(4)
Effect on Native Worker:				
- HC Spillovers of Native Co-Worker ($\overline{\alpha}_{-it}^{NN}$)	0.0901	0.1640	0.7060	0.8552
	(0.0234)			(0.0094)
- HC Spillovers of Immigrant Co-Worker ($\overline{\alpha}_{-it}^{NI}$)	-0.0501	0.0947	1.1665	1.3455
	(0.0337)			(0.0127)
Effect on Immigrant Worker:				
- HC Spillovers of Native Co-Worker ($\overline{\alpha}_{-it}^{IN}$)	0.0404	0.1578	0.9721	1.1930
	(0.0329)			(0.0152)
- HC Spillovers of Immigrant Co-Worker ($\overline{\alpha}_{-it}^{II}$)	0.0763	0.1725	0.7668	0.8989
	(0.0212)			(0.0067)
Worker Effects (α_i)	✓	✓	✓	✓
Year Effects (τ_t)			✓	
Establishment/Year Effects		✓		
Occupation/Year Effects		✓		
Firm/Occupation Effects			✓	
Establishment/Occupation/Year Effects (π_{eot})	✓			
Observations	1,012,748	1,012,748	1,012,748	1,012,748

Notes: The table report the estimates for the full-specification model using alternative fixed effect specifications. All specifications are controlled for age, age squared, tenure, tenure squared, gender, schooling, and nationality. Column (1) refers to the estimation results of regression (2), also presented in Table 2, Column (2). Columns (2)-(4) present alternative specifications in terms of different levels of fixed effects identified below the estimates. In Column (3), the alternative fixed effect specification uses firm instead of establishment due to the change of the establishment code at a certain year in the data set. Specification (1) presents the lowest estimates of all specifications as it controls most extensively for the different sources of sorting. In the other hand, specification (4) displays the highest estimates as it controls only for homophily through worker fixed effects. Due to computational power limitations, the standard errors for the specifications in Columns (2) and (3) were not estimated as the remaining specifications. Standard errors in parentheses and clustered at the peer group level.
Source: Quadros de Pessoal, 2002-2022.