2ND STUDY CYCLE
ECONOMICS AND BUSINESS ADMINISTRATION

Immigrant experiences in Portugal's private labor market: analyzing wage disparities

Sabrina Montenegro Chaves

2023

FACULDADE DE **ECONOMIA**

U.PORTO
FACULDADE DE ECONOMIA
UNIVERSIDADE DO PORTO

IMMIGRANT EXPERIENCES IN PORTUGAL'S PRIVATE LABOR MARKET: ANALYZING WAGE DISPARITIES

Sabrina Montenegro Chaves

Dissertation

Master in Economics and Business Administration

Supervised by
Vitor Miguel de Sousa Ribeiro

2023

Acknowledgments

Upon completing this academic stage, I couldn't help but express my profound gratitude to the people who made this master's degree a reality. Each contribution was essential to the success of this journey, making this achievement a shared accomplishment with everyone involved.

First and foremost, my heartfelt recognition goes to my supervisor, Professor Vitor Miguel Ribeiro, whose dedicated guidance, and profound knowledge, combined with his willingness to help, were crucial to the realization of this work. His incentive and inspiring vision were fundamental in conducting this study.

To my parents and brother, who are my unwavering support, I am grateful for the love, assistance, and understanding demonstrated every moment. Your constant encouragement during this journey was my greatest support, and without your backing, none of this would have been possible.

To my grandmother, who even from a distance supported and accompanied me exceptionally for the past two years.

To my friends and classmates, thank you for sharing this journey with me and making me feel welcome in this new country. Our discussions, exchange of ideas, and mutual learning made this academic experience much richer and more enjoyable.

I also want to thank the School of Economics and Management of the University of Porto for providing a valuable learning environment and opening doors to new opportunities. I extend these thanks to the professors whose knowledge-sharing enriched my academic years.

Finally, I would like to express my deep gratitude to my grandfather, Luis Montenegro Chaves, who is no longer here, but whose love and inspiration for our family continue to guide me in my life, becoming an eternal presence in my heart. Thank you for everything.

Thank you all for making this goal a reality.

Abstract

This study, through a matched employer–employee dataset, examined the wage disparity between immigrants and natives in the Portuguese private labor market for the period 2011-2020. By progressively incorporating new variables into the regression, applying unit and time fixed effects, employing Gelbach's decomposition technique, and utilizing a variable selection model, the persistence of wage divergence among nationalities in Portugal was confirmed. Such difference was found to be sensitive to the set of regressors used in the analysis, particularly highlighting the predictors tenure and qualification levels. Furthermore, the LASSO method indicated the maintenance of selected variables without a significant dimensionality reduction. Finally, the heterogeneity of the Portuguese market is emphasized, with significant differences among immigrant nationalities concerning the transferability of their qualifications to the market, the recognition of their skills by employers, and their remunerations.

Keywords: Wage gap, Labor market, Immigration, Portugal, Quadros de Pessoal.

JEL Classification: J15, J31, J61, J82.

Resumo

Este trabalho, através de um conjunto de dados longitudinais para os pares trabalhador-empresa, estudou a discrepância salarial entre imigrantes e nativos no mercado de trabalho privado Português para o período 2011-2020. Através da incorporação progressiva de novas variáveis à regressão, da aplicação de efeitos fixos temporais e individuais, da utilização da técnica de decomposição de Gelbach (2016), e da aplicação de um modelo de seleção de variáveis, foi confirmada a persistência da divergência salarial entre nacionalidades em Portugal. Tal diferença se revelou sensível ao conjunto de regressores utilizados na análise, com destaque para as preditoras antiguidade e níveis de qualificação. Ademais, a utilização do método LASSO, indicou a manutenção das variáveis aplicadas, sem uma significativa redução da dimensionalidade. Por fim, ressalta-se a heterogeneidade do mercado Português, com diferenças importantes entre as nacionalidades imigrantes em relação a transferibilidade de suas qualificações para o país, o reconhecimento de suas habilitações pelos empregadores e, consequentemente, nas suas remunerações.

Palavras-chave: Discrepância salarial, Mercado de trabalho, Imigração, Portugal, Quadros de Pessoal.

Códigos JEL: J15, J31, J61, J82.

Contents

List of Figures

List of Tables

1 Introduction

Human migration and mobility are complex and age-old phenomena that affect society worldwide. The World Migration Report 2022 reveals that in 2020, nearly 281 million people lived outside their countries of birth, with Europe being the primary destination, hosting 87 million foreign residents. In this context, Portugal has traditionally a history of emigration, with a significant percentage of its population residing abroad. However, this reality has been changing, and in 2019, the country reached unprecedented levels of immigrants, totaling over half a million (IOM, 2021; Oliveira, 2021).

Regarding the contribution of the foreign population to the Portuguese economy, it's noteworthy that in 2020, more than a billion euros destined for Social Security came from immigrants, demonstrating a higher contributory capacity than nationals. Furthermore, Oliveira (2021) also emphasizes the significant roles that foreigners assume in optimizing the efficiency of the labor market, bolstering the active workforce, mitigating the effects of the demographic aging of the population, and developing entrepreneurial initiatives. Consequently, the development of immigration policies that ensure the well-being of the foreign population and guarantee equity in the labor market is an increasingly relevant subject for the economic development of nations. Nevertheless, despite these contributions, an important area of the literature focuses on discrimination against immigrants, particularly concerning wages.

For instance, in the relatively recent landscape of studies on economic migration, pioneering authors like Chiswick (1978) and Borjas (1982) excel in research on immigrants' labor market adjustment and the reinforcement of the positive assimilation model. This theory posits that human capital is not perfectly transferable between countries and, therefore, upon arrival, foreigners present an average salary lower than nationals; however, after some years of migration and the development of specific skills from the destination country, immigrants gradually attain a level of education, qualification, and income comparable to that of local workers.

In the Portuguese literature, Carneiro *et al.* (2012) stand out for their prominent work on wage disparity, which contributes to the field of economic assimilation. Through their research, they examined matched employer-employee data and estimated Pooled Ordinary Least Squares (OLS) wage regressions, incorporating a dummy variable ("Immigrant") to signify foreign status. The authors progressively introduced new covariates related to workers' and employers' characteristics and subsequently incorporated establishments' fixed effects. Additionally, to assess the individual contribution of each variable in explaining the variation of "Immigrant" across different specifications, Gelbach's decomposition method was employed. The study's findings concluded that the earnings disparity between nationalities in the Portuguese private market is primarily attributed to the absence of match-specific human capital and occupational downgrading.

In this sense, this master's thesis contributes to the literature on wage disparity between immigrants and natives, following the study by Carneiro *et al.* (2012) in the Portuguese private labor market. However, it focuses on a more recent analysis period, from 2011 to 2020, and introduces additional individual fixed effects and an analysis of high dimensionality and variable selection. Furthermore, aiming to study the heterogeneity of foreigners in Portuguese territory, the immigrant population was segmented into different nationality groups, including those who arrived more recently in the country and thus were not previously considered in the literature, such as South Asians.

Additionally, it is mentioned that the chosen analysis period allows for the examination of the Portuguese market in an updated manner and also considers the effects of the 2010 financial and economic crisis, as well as the recovery periods following the intervention of the Troika.

Moreover, in accordance with other studies and considering the possibility of obtaining information with a high degree of reliability regarding the Portuguese private labor market, the Quadros de Pessoal was chosen as the data source. This longitudinal database allows for the control of the attributes of the workers (including nationality), of the companies, and of the match characteristics individuals form with firms, which makes it possible to identify the variables responsible for the wage disparity between immigrants and natives.

Finally, the paper is structured into four sections, apart from the introduction. Section two encompasses the literature review, presenting migration economics from a theoretical perspective. Furthermore, it proceeds with an analysis of international migrations, discussing trends and sociodemographic characteristics of the migrant population, with a particular focus on the Portuguese market. In section three, the construction of the study's sample is explained, along with a description of the variables of interest's means. Section four highlights the methodologies used in the research's development, presenting the results of these processes' application. Lastly, section five concludes the study with final considerations.

2 Literature review and migration trends

International human mobility and underlying economic factors, such as the wage gap between immigrants and natives, have garnered significant interest and debate in academic literature. Understanding the reasons for this disparity is essential for the formulation of policies and programs aimed at fostering the successful integration of immigrants into their host countries while also advancing equality and fairness within the labor market.

Consequently, this chapter offers a comprehensive overview of the relevant literature concerning the assimilation process of foreign workers, analyzing economic, social, and cultural factors that can contribute to wage discrepancies between natives and immigrants. Furthermore, it contextualizes major global trends in international migration and narrows the discussion to the Portuguese market.

2.1 Migrant literature

Despite migration being a recurrent phenomenon of human behavior, empirical evidence on the subject is relatively recent. Scholten *et al.* (2022) note that research in this area only began to formalize and expand in the 1950s and 1960s, focusing on sociological studies, such as the investigation of Gordon (1964), which addressed the integration of foreigners in host societies. However, Levy *et al.* (2020) explain that while sociology was the dominant discipline in early migration studies, the economic literature on the subject began to emerge more prominently in the 1970s.

During this time, the work of Chiswick (1978) was relevant for research on assimilation and the economic impact of immigrants on the labor market. The study examined cross-sectional data in the United States by formulating a linear regression of the natural logarithm of annual earnings (wages, salary, and self-employment income) on the exogenous variables: years completed of schooling, labor market experience, years since migration, dichotomous variable equal to unity for a person of foreign birth, marital status, and other relevant terms. The findings revealed that upon arrival in the country, immigrant men had significantly lower earnings than native workers. However, foreigners experienced relatively rapid wage growth that eventually peaked and surpassed local workers. Chiswick (1978) estimated that this turning point occurred around 10 to 15 years after migration, with immigrant qualifications, the transferability, and the origin of their human capital influencing the occupational trajectory. These results were groundbreaking in introducing the concept of human capital to explain the empirical evidence and justify the economic incentive for immigrants to acquire specific qualifications in their destination country (Borjas, 1985).

Building upon Chiswick's research, subsequent studies delved into analyzing wage gaps, often focusing their investigations on different ethnic groups, databases, and gender. For instance, Borjas (1982), in an analysis of immigrants of Hispanic origin in the U.S. labor market, observed a positive correlation between the group's income and the years since migration, with the rate of wage response to assimilation varying significantly between different Hispanic nationalities. Notably, the author emphasized that investment in human capital could account for the superior positive performance of certain immigrant groups in the U.S., generating faster rates of economic mobility.

Therefore, in line with various studies after Chiswick (1978), the results of Borjas (1982) tend to confirm consistent patterns in the U.S. market: an initial disadvantage for immigrants followed by recovery within a span of less than two decades of assimilation. However, Borjas (1985) questions the empirical validity of this conclusion, asserting that the cross-section regressions commonly employed in the literature do not adequately account for changes in the characteristics of immigrant cohorts over time. In other words, relying on different observations at a single point in time does not allow for the monitoring of how immigrants' human capital and living conditions evolved over the years. Consequently, the true assimilatory effect may not be accurately captured.

Thus, when examining remunerations within immigrant cohorts, Borjas (1985) found that most foreign groups in the U.S. market experienced relatively slower earnings growth. He concluded that the cross-section regressions had overestimated the rate of increase experienced by foreigners by as much as 20 percentage points. Cabral & Duarte (2013) mention that one alternative to this problem involves employing longitudinal databases that track both native and immigrant populations over time, thereby enabling the estimation of the assimilation trajectory for immigrants who have resided in the country for a specified duration.

Regarding the investment in human capital, as previously mentioned, the literature on immigrant earnings has explored whether acquiring qualifications in the destination country leads to wage growth for foreign-born individuals. For instance, Duleep & Regets (1999) concluded that immigrants with limited initial transferability of qualifications, especially those from underdeveloped countries, exhibited a substantial propensity for investing in skills and, therefore, to experience growth in earnings. The authors showed that this group was more likely to invest in human capital development than both U.S. workers and immigrants with greater ease of transferring skills. As a result, these newly arrived foreigners were more adept at adjusting to the evolving skill requirements of the economy than native-born individuals and highly skilled immigrants.

Similarly, Longva & Raaum (2003) corroborated with the previous literature findings, observing that upon their arrival in Norway, immigrants from non-OECD nations experienced notably lower wage returns than natives. However, the remunerations gradually increased with the duration of their stay in the country and the accumulation of specific human capital. The authors highlighted that the transferability of qualifications varies significantly among foreigners, leading to diverse growth trajectories as the years since migration increase.

Finally, in a more recent study on Germany, Aldashev *et al.* (2012) showed that acquiring a German education was a fundamental component of reducing the significant initial immigrant wage disadvantage. However, the study also found that a significant portion of the wage gap remained unexplained even after accounting for education acquired in Germany. One possible reason is that educational degrees obtained abroad may not be comparable.

Therefore, the literature documents lower initial marginal returns on immigrant capital, highlighting the imperfect portability of qualifications and the need to acquire specific skills in the destination country as the reason. Sanromá *et al.* (2015) explain that this reduced valuation of immigrant qualifications on arrival is due to three main issues. First, insufficient command of the language spoken in the host country restricts the immigrant's ability to contribute and can limit his sense of belonging. Similarly, the academic training offered in the origin country may suffer from insufficiencies that translate into a lower quality of education. Furthermore, cultural, social, institutional, technological, and economic aspects of the nation of precedence are specific factors considered when developing the immigrant's professional experience, often making them inexperienced in the destination and resulting in lower productivity.

However, it is worth mentioning that although productive differences between immigrants and natives are important, the occupational mobility of foreigners is also an aspect cited in the literature as relevant to understanding the wage gap upon arrival in the destination country. Damas de Matos (2017) explains that newly arrived immigrants may have a higher propensity to accept job offers with lower salaries, primarily due to their limited knowledge about the labor market. On the other hand, companies that pay better may not agree to hire individuals without previous experience in the country. Thus, immigrants suffer from an initial downward trajectory associated with occupying less qualified positions than those they used to assume in their origin country. Likewise, Carneiro *et al.* (2012) also show the sensitivity of the wage gap to qualification levels, emphasizing that immigrants work in lower positions on the occupational scale than natives with similar productive characteristics.

Furthermore, Carneiro *et al.* (2012) explain that, besides occupational downgrading and the lack of specific human capital in the destination country, the concentration of foreigners in workplaces is also a relevant factor in explaining wage gaps. Thus, the authors show that immigrants are attracted to companies with a higher proportion of non-national employees, causing a reduction of 0.19% in the remuneration of foreign men for each 1% increase in concentration. According to the authors, this penalty can be explained, for example, by the greater willingness of immigrant workers to work with other foreigners, especially if they have cultural proximity such as language and work habits, being, because of this, more willing to accept lower wages.

In addition to the factors already mentioned, the literature also considers the wage discrepancy in light of the discriminatory aspect. Oaxaca (1973) and Blinder (1973) stand out in this field by proposing the wage decomposition observed in the economic results of two groups, such as men and women or immigrants and natives. This methodology seeks to understand how much of the observed difference can be explained by differences in individuals' characteristics, such as education, experience, and economic activity, and when it is due to discrimination factors, such as gender stereotypes and xenophobia.

According to Nielsen *et al.* (2004), three theoretical approaches to discrimination seek to explain the existence of wage discrepancies between equally productive workers. The first is related to the concept of "taste-based discrimination" proposed by Becker (1957), in which employers, co-workers, or customers assume unfavorable attitudes towards ethnic or gender minorities, resulting in a lower probability of hiring such a group and in a reduced willingness to pay wages commensurate with the qualifications of these individuals.

The second theoretical group argues that employers may not have accurate information about the future productivity of candidates, which leads them to make inferences based on observable characteristics, such as gender or ethnicity, to estimate the future performance of workers. This scenario is known as statistical discrimination, in which employers base hiring decisions, job allocation, or wage setting on the worker's association with certain demographic groups (Tilcsik, 2021).

Finally, the third approach concerns the labor market division into sectors with different characteristics. These segments can be defined according to economic activity, type of employment contract, educational level, and other factors, which result in different conditions and employment opportunities for workers. For example, if immigrants tend to choose qualifications or skills associated with lower-paying occupations, this labor market segmentation can result in wage differences compared to native workers (Nielsen *et al.*, 2004).

Additionally, it's worth noting that early research on migration focused primarily on the United States market, as shown in the previously mentioned studies: Chiswick (1978), Borjas (1982, 1985), and Duleep & Regets (1999). However, with the development of the literature and the progressive increase of the migratory flow directed to the countries of southern Europe at the beginning of the 21st century, studies appeared on the Iberian Peninsula market, especially Portugal and Spain.

In this context, in Portuguese literature, the research by Carneiro *et al.* (2012) for 2003-2008 estimated three wage equations using the Pooled OLS method, progressively incorporating new covariates related to the workers' and employers' characteristics and, subsequently, considering establishments fixed effects. The sample was also divided by gender for differentiated analysis.

Similarly, Cabral & Duarte (2012, 2016) also carried out studies using the Ordinary Least Squares method, estimating wage regressions considering the main nationality groups of immigrants in Portugal. Finally, Damas de Matos (2017) contributed to understanding the wage gap between foreigners and natives in the Portuguese market using data from 2002 to 2009. The study employed logarithmic wage regressions considering workers' and firms' fixed effects.

The mentioned studies highlight that there are no significant contributions in the Portuguese literature, apart from those of Carneiro *et al.* (2012), to understand the impact of gender in the study of the immigrant wage gap. In this regard, it is worth mentioning studies such as Boyd (1984) and Beach & Worswick (1993) in the Canadian market, Card (2005) in the United States, and Nicodemo & Ramos (2012) in Spain, which sought to assess whether gender adds another dimension to the stratification of immigrants in the labor market. The results of these surveys can be summarized as follows: first, some groups of immigrant women earn, on average, less than the native population, and second, less than immigrant men. Nicodemo & Ramos (2012) describe this phenomenon as a double wage penalty.

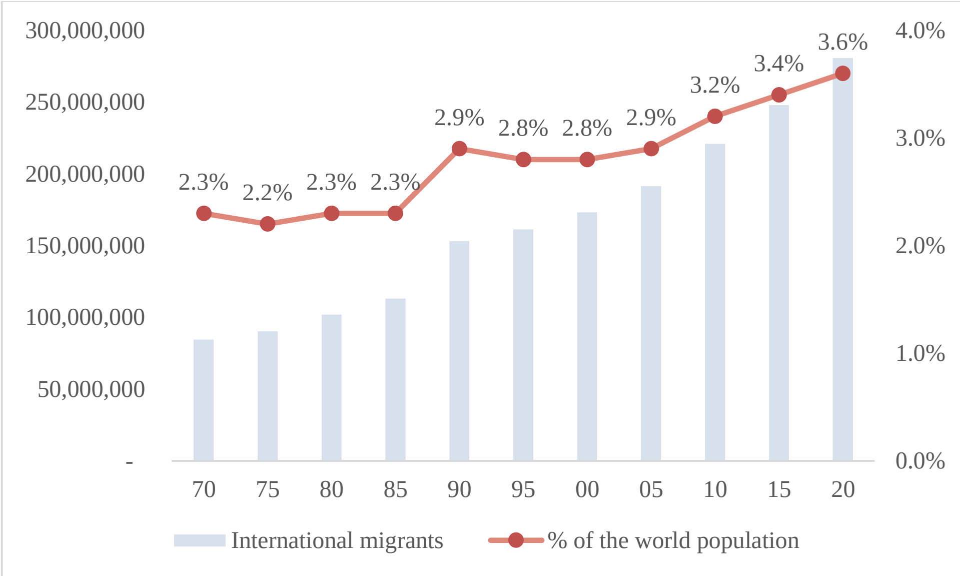
Boyd (1984) reinforces this idea, explaining that individuals often face disadvantages for belonging to groups that are negatively evaluated in statistics based on criteria of gender, country of birth, and race and that participation in one or more of these groups aggravates these disadvantages. Therefore, the double discrimination of immigrants is related to their female and foreign status. However, this characteristic tends to be less significant for immigrants from developed countries (Nicodemo & Ramos, 2012).

Long (1980), contradicting the results found in the literature, concluded, when studying the U.S. market, that the earnings of foreign women were approximately 13% higher than those of native women. On the other hand, Nicodemo & Ramos (2012) found that immigrants received less than women born in Spain and that characteristics present in natives and absent in foreigners were better rewarded in the labor market.

2.2 International migration trends

According to the World Migration Report 2022, an estimated 281 million international migrants were recorded globally in 2020, which accounts for 3.6% of the global population and exceeds more than three times the number reported in 1970 (84 million). Between 2000 and 2010, the number of people residing outside their countries of origin increased by approximately 48 million, with an additional 60 million added to the estimates between 2010 and 2020. Although the growth seems substantial, it represents only a 1.3 percentage point increase compared to the world's population in 50 years (see Figure 1). Therefore, it is noteworthy that staying in the country of origin remains the predominant norm. According to the report, domestic mobility, movement within national boundaries, still stands out as a phenomenon of greater relevance than international migration (UN/DESA, 2020a; IOM, 2021).

Figure 1 - Number and share of international migrants (1970 - 2020)



Source: IOM. Own elaboration.

Furthermore, international migration is primarily driven by income and security-related issues involving voluntary and forced migrations. The United Nations High Commissioner for Refugees reported that, in 2019, the number of forcibly displaced people due to persecution, violence, or human rights violations reached 79.5 million, of which one-third were refugees. However, during the same year, approximately 169 million migrant workers were registered globally, representing a significant portion of the labor force in many countries and making relevant contributions to economies and society. Nevertheless, migrant workers are often found in temporary and informal jobs and are subjected to unequal treatment in the labor market (ILO, 2020; 2021; IOM, 2021; UNHCR, 2023a).

Of the working migrant population, 96% resided in medium and high-income countries in 2019. The proportion decreased by 7.3 percentage points between 2013 and 2019 in high-income nations, while there was an increase of 7.8 percentage points in medium-income countries during the same period. This variation may indicate either economic growth in medium-income countries or possible changes in labor immigration regulations in high-income nations (IOM, 2021).

In this context, Europe was the principal destination for migration in 2020, with 30.9% of the international migrant population, followed by Asia with 30.5%. However, Oceania stood out as the continent with the highest proportion of migrants relative to its total population, with 22% of its residents born abroad. It is worth noting that this scenario has persisted since the 1990s, with Europe and Asia consistently occupying the top positions (IOM, 2021).

Regarding nations, the United States remained the largest immigrant-receiving country in 2020, with 51 million foreign-born individuals. Germany, the United Arab Emirates, the Russian Federation, and the United Kingdom follow, comprising the group of the top five destinations for migratory flows. From Oceania, Australia stands out, ranking ninth on the list (IOM, 2021).

Concerning countries of origin, India had the highest number of emigrants, with approximately 18 million citizens. Mexico ranked second with 11 million individuals residing in other nations, closely followed by Russia and China, with around 10.8 and 10 million emigrants in 2020, respectively. The Syrian Arab Republic completed the group of the top five origins, particularly notable for its refugee population due to the country's humanitarian crisis (IOM, 2021).

Moreover, the proportion of male migrants is higher than that of females, with 52.1% of the group in 2020. Such a trend shows that, despite men being more likely to migrate, women constitute almost half of the international migrant population. This result is interesting as, for years, the literature has considered migration as a predominantly male phenomenon. In this context, Lutz (2010) explains that initially, studies aimed only to demonstrate patterns of female migration without significantly changing the perception of the immigrant profile, persisting in the association of women as dependents on their husbands or other family members. A more prominent interest in the literature regarding female migrants began with the development of Women's Studies in the 1980s, which highlighted gender relations as determinants of migration. That is the set of social representations, roles, and behaviors of women and men as explanatory factors for migratory reasons (Boyd & Grieco, 2003; Lutz, 2010; IOM, 2021).

In this context, there is currently a notable increase in the number of women migrating alone for study and work purposes. Many of them have become the principal financial providers for their families, sending remittances back to their countries of origin and thus contributing to their communities' economies. Despite this, migrant women remain the most vulnerable group, facing risks of exploitation and violence. According to the 2022 Global Report on Trafficking in Persons by UNODC, women and girls accounted for 60% of the total number of detected trafficking victims in 2020. However, countries have been making strides in increasing the implementation of formal mechanisms to ensure that migration policy is gender-sensitive, including advisory committees and similar bodies. Thus, a study published in the 2019 World Population Policies Report by UN/DESA documented that 65% of governments in Europe and North America have mechanisms in place to ensure that migration policies are gender-sensitive (UN/ DESA, 2020a; 2020b; UNODC, 2022).

Regarding the volume of financial remittances, there is a growing trend in international money transfers, increasing from US$ 128 billion in 2000 to US$ 702 billion in 2020, with countries such as India, China, and Mexico standing out as the leading recipients. These remittances sent by migrants help improve the quality of life for their families in their home countries, contributing to education, health, housing, and sanitation (UN/ DESA, 2020a; IOM, 2021).

It is also noteworthy that 78% of international migrants were within the age group of 15 to 64 years old in 2020. Nevertheless, the proportion of individuals under 19 years old has decreased since 1990, reaching 14.6% of the migrant population in 2020. This indicates a negative difference of 4.3 percentage points over the past 30 years. Consequently, it becomes evident that most migrants belong to the active working-age group (IOM, 2021).

Among the economic activities with the highest concentrations of international migrants, a notable observation is that in 2019, the services sector employed 66.2% of the total workforce and approximately 79.9% of female migrants. The increasing demand for professionals in the care economy (including domestic and human health activities), where most workers are women, may explain the high percentage of female migrants in the services sector compared to men, whose proportion is 56.4%. For male migrants, the focus is on manufacturing industries, with increased demand in low and middle-income countries. In addition to services and manufacturing, agriculture covers 7.1% of the total migrant workforce (ILO, 2021).

Regarding the motivations that drive individuals to migrate, opportunities for positive advancement, often associated with financial reasons, are strongly linked to international migration. However, the desire to migrate does not necessarily translate into migratory action, as the feasibility of this decision is intricately related to the infrastructure involved in the process. Aspects such as regulatory regimes and political structures play fundamental roles in enabling or restricting the realization of the act. However, it is worth noting that, in addition to political factors, international migration presents additional challenges and complexities that determine the economic achievements of migrants (ILO, 2020; IOM, 2021).

As previously mentioned, linguistic skills and other elements related to human capital, along with the transferability of these aspects to the new market, are relevant factors that determine individuals' success in their destinations. According to the International Labour Organization, the skills of migrants are often not fully recognized, primarily due to the lack of appropriate recognition systems. As a result, migrants often find themselves compelled to seek lower-skilled positions that do not match their educational qualifications or professional experiences (ILO, 2020; IOM, 2021).

The role of education in migration decisions is also noteworthy. According to the 2019 Global Education Monitoring Report by UNESCO, individuals with primary education are twice as likely to migrate, those with secondary education are three times more likely, and those with university education are four times more likely compared to individuals without formal qualifications (UNESCO, 2018). Furthermore, Goldin et al. (2018) mention that between 1999 and 2010, the number of international migrants with tertiary education increased by 130%, which increased highly skilled workers in the labor markets.

Finally, concerning the contribution of migration to host countries, Basso & Peri (2020) explain that the presence of migrants results in an increase in the number of active workers due to their younger age profile compared to the native populations. This factor drives the growth of these countries' economic production and, consequently, results in higher per capita and aggregate GDP. Additionally, the proportion of workers relative to dependents in the economies increases, meaning there are more individuals of working age to support the portion of the population that does not work, including children and seniors (Goldin *et al.*, 2018; Basso & Peri, 2020; IOM, 2021).

Additionally, it is worth noting that migrants tend to exhibit greater flexibility regarding changes in location and sectors compared to native residents, possibly due to less pronounced family and social ties, contributing to increased labor market efficiency. Moreover, there is also evidence regarding the impulses for innovation generated by skilled immigration. In the United States, for example, Hunt & Gauthier-Loiselle (2010) found that a 0.45 percentage point increase in immigrant scientists and engineers contributed to approximately a 13% rise in patents per capita between 1990 and 2000.

Lastly, Csonto *et al.* (2015) mention that the net fiscal impact of immigration may initially be negative due to the cost of humanitarian support for refugees and integration policies. However, the effect becomes positive as the workforce integrates migrants. Thus, the authors conclude that, in the long term, the advantages of migration are significant, creating opportunities in terms of remittances, trade, and investment flows.

However, the literature does not consistently reinforce this positive portrayal of migration. For example, Borjas (2003) argues that immigration reduces the wages of competing workers, as a ten percent increase in the labor market supply in the United States leads to a three to four percent wage reduction. In this regard, the author contends that immigrants can cause wealth redistribution in the economy and reduce income by increasing competition in the labor market.

Nevertheless, De Brauw (2017) maintains that the negative wage effect would only be evident if immigrants and natives were perfect substitutes and, therefore, competed for the same jobs. He explains that typically, the two groups accept different job opportunities since they possess distinct comparative advantages. However, the author confirms that the divergences in empirical evidence on the topic are due to varying methods used by economists to study the effects of migration.

It is important to highlight that despite advocating for wage reduction in the economy, Borjas (2003) reinforces that the influx of immigrants can potentially be a net benefit for the nation, increasing the total wealth of the population.

2.3 Migration in Portugal

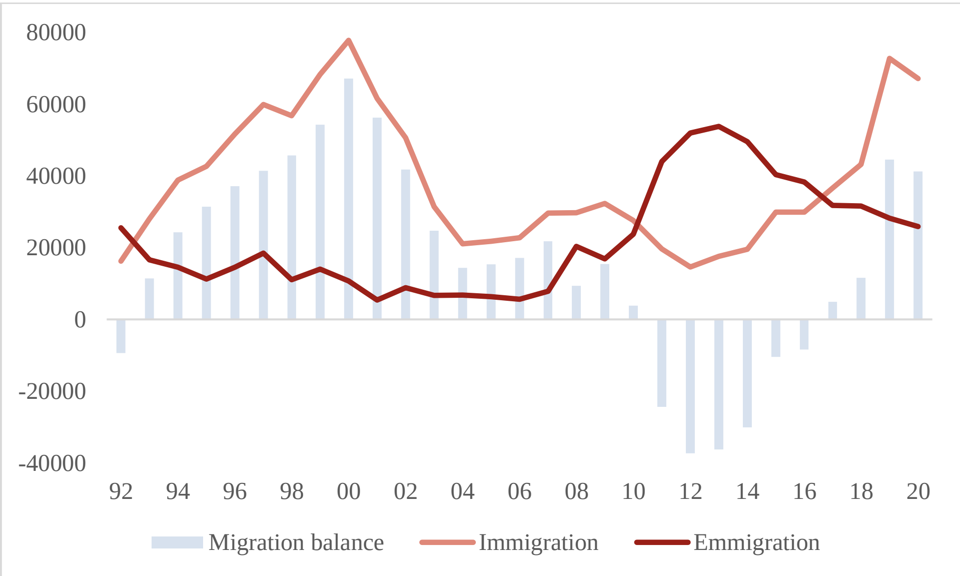
In Portugal, Góis *et al.* (2018) explain that migratory flows represent a complex and multidimensional reality, including seasonal migration driven by tourism, temporary immigration of students, and long-term migration involving labor and family reunifications. However, the relationship between the country and migratory phenomena has not historically been consistent. Cabral & Duarte (2012) explain that until the 1990s, immigration to Portuguese territory was moderate, mainly driven by labor reasons. Despite significant changes in this reality, with a widespread increase in the number of immigrants, Oliveira (2021) reveals that Portugal diverged from most European countries by experiencing negative net migration between 2011 and 2016, a situation associated with the economic crisis and the native population's search for better financial opportunities.

In this way, over the past four decades, Portugal has experienced significant transformations concerning its role in international migrations. The country actively participated in the main migratory flows of the 20th century, with intense emigration initially directed to Brazil and Venezuela and later to the United States and Canada. In the case of Brazil, the volume and constancy of Portuguese immigrants were primarily influenced by the historical colonization process. Regarding Venezuela, in the 1940s, there was a significant influx of Portuguese immigrants into the country due to the difficult living conditions in post-war Europe and the favorable immigration policies implemented by the Venezuelan government (Gomes, 2009; Malheiros & Esteves, 2013).

Starting from the 1950s, European countries such as Germany and France began to attract Portuguese workers who sought better economic opportunities and wished to leave the rural environment of Portugal. It is relevant to note that until this period (mid-last century), foreigners in the country comprised a small portion of the population (around 30 thousand foreigners in 1960) and were mainly from Portuguese overseas territories, as well as nations that invested in the country, such as Spain, the United Kingdom, and France (Malheiros & Esteves, 2013).

With the end of the Salazar dictatorship and the independence of the Portuguese colonies in Africa in the 1970s, the number of immigrants in Portugal started to rise, reaching 85 thousand foreign individuals in 1985, with a notable presence of the Cape Verdean population. However, with Portugal's accession to the European Community in 1986 and the subsequent economic stimulus, the entry of foreign residents accelerated.

Figure 2 - Migration flows and balance in Portugal (1992 - 2020)



Source: Eurostat. Own elaboration.

Consequently, as employment opportunities increased, the country witnessed positive migratory flows for the first time in the early 1990s, as observed in Figure 2. Thus, among the 27 European Union (EU) member states, Portugal held the sixth position in 2020 in terms of its migratory balance, ranking behind only Spain, Germany, the Netherlands, Belgium, and France (Malheiros & Esteves, 2013; Oliveira, 2021).

In this way, there was an increase in the foreign population until the beginning of the economic and financial crisis in 2010, which led to a substantial reduction in the issuance of new residence permits. However, with the easing of the recession from 2016 onwards, new growth was observed, mainly driven by the granting of visas to retirees (which doubled from 2016 to 2017), investors (+59% from 2017 to 2018), highly skilled workers (+65.7% from 2018 to 2019), and students (+47% from 2016 to 2017) (Malheiros & Esteves, 2013; Oliveira, 2021).

Therefore, in 2020, the number of immigrants with valid residence permits in Portugal reached 661,607, accounting for 6.4% of the total population. This value represents an increase of 4.4 percentage points in the number of foreigners relative to the total residents since 2000. However, it is important to highlight that, despite the sustained growth of the foreign community, Portugal has not ceased to be an emigrant nation and recorded 25,886 permanent emigrants in 2020 (Oliveira, 2021).

It is worth mentioning that between 1992 and 2007, the Portuguese government approved six processes for immigrants' regularization, providing statistical visibility of the foreign population. However, the legal framework in Portugal in 2006, which brought about the fourth amendment to Law No. 37/81 (Nationality Law), expanded immigrants' rights to acquire Portuguese nationality. Through the approval of dual nationality and the reduced importance of blood relation or filiation, 223,231 nationality requests were granted between 2007 and 2012, leading to the statistical disappearance of a significant portion of foreign citizens (Oliveira & Gomes, 2014; Sampaio, 2017).

Furthermore, it is relevant to emphasize that the implementation of immigration policies played a fundamental role in encouraging the entry of foreign residents. Pinho (2013) explains that until the 1980s, the existing legal framework was fragmented and mainly focused on regulating the expulsion of foreigners and the right to asylum. It was in 1981, with the emergence of the Nationality Law, which defined the conditions for access to Portuguese citizenship, that the first comprehensive set of laws to regulate the entry, stay, and exit of immigrants from the national territory was established. Subsequently, in

1998, new legislation was implemented to combat irregular immigration while ensuring rights to foreigners related to family reunification. In 2003, an annual maximum limit for the entry of third-country nationals was enacted based on the labor needs of each sector of the Portuguese economy. The approval of this law established, for the first time, a connection between the admission of immigrants and national interests (Pinho, 2013).

Currently, the legal framework for immigration, notably the Law on Foreign Nationals, establishes all procedures and requirements for visa issuance while ensuring the rights and integration of those who choose to live and work in Portugal. Furthermore, it aims to facilitate the entry of higher education students and citizens from the Community of Portuguese-Speaking Countries (CPLP) (ACM, 2022).

In addition to applicable legislation, various immigration incentive programs have been developed over the years to attract, primarily, highly skilled labor. In this context, in 2009, Portugal established the "Non-Habitual Residents Tax Regime", aiming to attract highly qualified foreign professionals with significant assets to the country. Under this program, for ten years, earnings obtained within the national territory are subject to a flat tax rate of 20%, while incomes acquired outside Portugal are tax-exempt (AT, 2016).

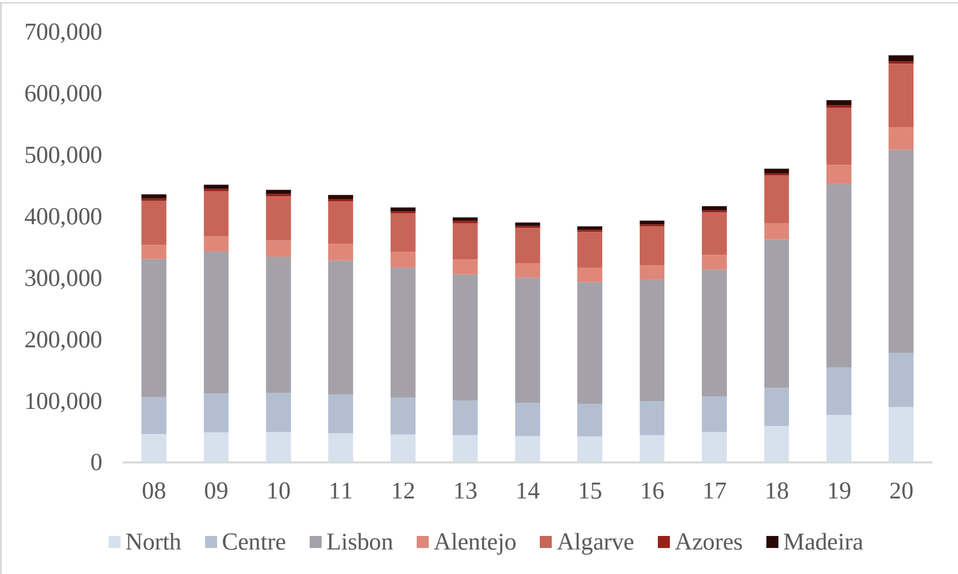
In 2012, the "Golden Visa" project was established to simplify the regularization process for foreign investors in Portugal, including property acquisition or capital transfer. It is worth noting that in 2014, 4.1% of the total private investment in the country came from Golden Visas, demonstrating the program's contribution to the Portuguese economy. Also, aiming to attract investors, the "Startup Visa" was launched in 2018, targeting entrepreneurs without permanent residence in the Schengen Area who wish to develop innovative projects in the country. In addition to expediting visa acquisition for these professionals, the program offers support to the projects through a network of incubators provided by the Agency for Competitiveness and Innovation (IAPMEI) (SEF, 2012; IAPMEI, 2022).

Finally, it is worth mentioning the "Tech Visa", aimed at facilitating the access of high-level professionals in the technology sector to job opportunities in Portugal, and the "Interior Employment PLUS" program, created to support the geographical mobility of professionals to regions in the interior of the country, thus encouraging the economic and social development of these areas (Portugal, 2018, 2020).

Regarding the sociodemographic characteristics of the immigrant population, Oliveira (2021) explains that, in 2019, there was a reversal of the trend observed since 2012 regarding the feminization of the immigrant population. As a result, men surpassed women again, and in 2020, they represented 50.8% of the total immigrants in the country. Despite this, similar to the international migration trend, female immigration in Portugal is no longer predominantly associated with family plans, and there has been an increase in the number of women who move independently and autonomously. Currently, they stand out in the allocation of residence visas for studies, professional internships, and volunteer activities, accounting for 51.4% of this category in 2020 (Oliveira, 2021).

It is also noticeable that the foreign population is not distributed homogeneously throughout Portugal, with significant differences between the country's districts. Thus, in 2020, Lisbon, Faro, Setúbal, and Porto accommodated 75.5% of immigrants, while Vila Real and Guarda held the last positions, with 0.8% of the total foreign citizens. However, Viana do Castelo and Viseu were the districts with the highest positive variations between 2019 and 2020: 18.6% and 18.4%, respectively. In contrast, Azores (4.9%) and Vila Real (+6.1%) showed the lowest increases in the foreign population (Oliveira, 2021).

Figure 3 - Distribution of immigrants across Portuguese NUTS II regions (2008 - 2020)



Source: FFMS. Own elaboration.

Figure 3 shows the evolution of immigrants concerning the secondary subdivision of the Nomenclature of Territorial Units for Statistics (NUTS), confirming the heterogeneity of the distribution in the Portuguese territory, with Lisbon being the main center of concentration for the foreign population. Furthermore, the figure confirms the reduction mentioned before in the number of immigrants during the 2010 economic crisis, followed by a subsequent recovery in 2016.

Concerning the age structure, Oliveira (2021) states that the foreign population, as observed in international migration, tends to be younger than the Portuguese population, with a predominance of individuals in the working-age group. In 2019, approximately 60.6% of immigrants were between the ages of 20 and 49, while for native residents, the percentage decreased to 36.6%. On the other hand, 22.9% of Portuguese citizens were 65 years or older, contrasting with the value of 9.5% for foreigners. Additionally, it is worth mentioning that immigrants not only contribute to increasing the number of active workers in the country but also to boosting births in Portugal. In this sense, 12.7% of the births registered in the nation in 2019 were from immigrant mothers.

Furthermore, regarding the countries of origin, it is observed that in 2020, Brazil remained the main foreign nationality in Portuguese territory, accounting for 27.8% of the total immigrant population. The United Kingdom, Cape Verde, Romania, and Ukraine followed, comprising 21.3% of foreigners. The case of Eastern European countries is interesting because the entry of such nationalities occurred suddenly and intensely from the early 2000s: approximately 70,430 citizens from the group were granted residence permits in 2001 (a positive variation of 2,968% since 1999), surpassing all Lusophone nationalities. This migratory flow was unpredictable, as the Portuguese government did not adopt any policies to encourage the entry of these immigrants during that period, and there were no historical or economic ties between Portugal and these countries that justified such demand. Baganha *et al.* (2004) explain that this flow may be linked to migratory pressure in Eastern Europe in the late 1990s, which was directed toward countries like Portugal, where the economy was expanding and facing labor shortages.

Regarding the labor market, Góis *et al.* (2018) mention the bipolarity of the Portuguese employment sector in the late 1990s, revealing the segmentation faced by immigrants when performing either low or highly-skilled professions, with a lack of intermediate job positions available for this group. Thus, Oliveira (2021) notes the high demand among the immigrant population for precarious and lower-paid jobs, concentrated in sectors like construction, hospitality, and food and domestic services. However, there have been changes in recent decades, mainly driven by information and communication technologies, which not only facilitate migration and hiring processes but also create opportunities for highly skilled professionals in the sector. Additionally, there has been a significant increase, between 2011 and 2019, in the number of foreign workers in intellectual and scientific activities, corresponding to a positive variation of 89.1% (Góis & Marques, 2007; Oliveira, 2021).

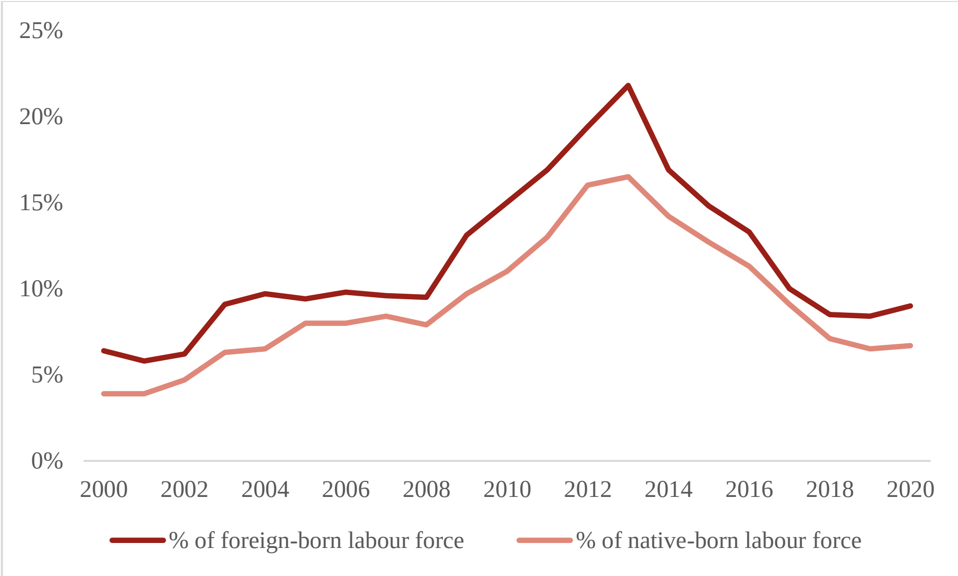
Moreover, Góis *et al.* (2019) highlight entrepreneurship as a relevant factor among the immigrant population in Portugal. According to the authors, the relative importance of foreign employers in the total national workforce has increased in recent years, rising from 4.2% in 2001 to 5.2% in 2011. These entrepreneurial initiatives contribute positively to immigrant integration, promote dynamism and innovation in the economy, and create new job opportunities for locals. Oliveira (2021) reinforces this, explaining that immigrants show a higher level of entrepreneurship than Portuguese nationals, which, in turn, plays a crucial role in job creation.

It is also relevant to highlight the contribution of immigrants to Portuguese Social Security. Oliveira (2021) explains that foreigners contribute more than Portuguese nationals, playing an essential role in sustaining Portugal's social security system. In 2019, this group contributed a significant amount of 884 million euros. On the other hand, immigrants still have fewer Social Security system beneficiaries than Portuguese nationals.

Concerning educational qualifications, there has been a significant increase in the number of international students enrolled in higher education in recent years, corresponding to a 160% increase between 2011 and 2019. Part of this growth is related to government programs aimed at attracting international students and, as mentioned earlier, the simplification of visa acquisition processes. Thus, Oliveira (2021) confirms an increase in the number of foreigners with medium-high educational qualifications and a decrease in individuals with primary and secondary qualifications, strengthening the Portuguese job market with qualified professionals.

However, until 2007, legislation for the recognition of foreign qualifications was still underdeveloped and limited to countries where Portugal had reciprocal agreements, making it difficult for immigrants to assume positions equivalent to their educational degrees. In 2018, Decree-Law no. 66/2018 was approved, introducing a new legal framework that facilitated and standardized recognition procedures, promoting progress toward the portability of human capital between immigrant countries and Portugal (Oliveira, 2021).

Figure 4 - Foreign and native-born unemployment in Portugal (2000 - 2020)



Source: OECD. Own elaboration.

Finally, it is worth highlighting the analysis of the unemployment rate among the foreign population. As mentioned by Oliveira (2021), the job precarity faced by immigrant workers puts them at greater vulnerability and increases their unemployment rates during periods of recession. Thus, observing Figure 4, it is evident that foreigners consistently exhibit a higher proportion than Portuguese nationals, with the difference increasing during the 2010 financial crisis, reaching a peak of 5.3 percentage points in 2013. Additionally, it is noticeable that the unemployment rate among immigrants rises more sharply during recessionary periods than among native workers. However, after the easing of the crisis, the gap between Portuguese and immigrant rates decreased significantly, reaching a minimum value of 0.9 percentage points in 2017. This scenario reflects that during periods of economic growth, foreign labor is highly valued but becomes dispensable in times of financial crisis.

3 Portuguese matched employer-employee data

This analysis aims to perform an econometric examination of the wage disparities between foreign-born and domestic workers in Portugal. Therefore, an all-encompassing and trustworthy database on the Portuguese labor market has been a crucial motivator for this research.

Hence, the following sections present an extensive depiction of this database, the Quadros de Pessoal (QP), and elaborate on the steps to establish the study sample. Furthermore, a descriptive analysis summarizes the primary characteristics and patterns of the variables of interest.

3.1 Data

Accurate and unbiased statistical information is necessary for economic agents and governments to make strategic decisions and respond effectively to current circumstances. In this regard, Portugal benefits from the publication of the Quadros de Pessoal, a longitudinal dataset for employer-employee pairs active in the country, conducted annually by the Office of Strategy and Planning (GEP) of the Ministry of Labor, Solidarity and Social Security (MTSSS) (Cabral & Duarte, 2012). The QP surveys are made mandatory and publicly disclosed to workers, covering all private sector firms with employees, and providing reliable and extensive information on the business structure, employment, duration, wages, and collective labor regulation (Carneiro *et al.*, 2012).

However, it is essential to note that the QP surveys do not include information on central, regional, and local administration, public institutes, and domestic workers' employers. Additionally, the data only indicate when employees started working in activities covered by the questionnaire, which may not correspond to their arrival in the country for non-national workers. Concerning educational level, the QP does not make it possible to ascertain the country where the worker acquired formal education, limiting the distinction between Portuguese and foreign qualifications. Since the end of the first decade of the 21st century, there has been a shift in the main reasons for immigrants' entry into Portugal, with a decrease in labor-related flows due to the country's economic situation and an increase in family reunification and study visas (Oliveira, 2021). Thus, it is unreasonable to assume that foreign workers obtained education in their country of origin, and the QP represents a limitation in assessing human capital portability.

It is also pertinent to mention that the data do not inform about the individual's country of birth, only their nationality. Therefore, there may be immigrant workers who acquired Portuguese nationality and started registering as nationals from the year of naturalization. Subsequently, Heckman's selection method will be applied to study this potential selectivity problem.

Despite these limitations, the QP surveys provide valuable insights into workers' characteristics, including gender, age, tenure, type of contract, professional category, educational qualification, nationality, level of qualification, regular and overtime monthly hours worked, and remuneration data, such as base and extra salary, and optional benefits, that is, gross amounts paid periodically or not periodically, like meal allowances and profit-sharing bonds. The database also offers information on firms, such as the year of incorporation, legal form, share capital, location, economic activity, turnover, and number of employees.

3.2 Sample definition

Concerning the period covered by the study, the statistical data for the QP has been compiled since 1981, although information regarding workers' nationalities was not available until 2000. Furthermore, some variables - including the Nomenclature of Territorial Units for Statistics, Portuguese Classification of Economic Activities (CAE), and educational qualifications - have undergone reclassifications over the years.

It is also noteworthy that there is an interest in evaluating the effect of the economic and financial crisis on immigration in Portugal. As Mota (2017) mentions, in 2010, due to the U.S. subprime crisis, many economies were in a state of recession or stagnation, with high unemployment and indebtedness. One consequence was the triggering of the sovereign debt crisis in the peripheral countries of the Eurozone, including Portugal, leading to expectations of default on public debt induced by high interest rates. As a result of this recession, the Portuguese government requested assistance from the Troika, a group composed of the International Monetary Fund (IMF), European Commission, and European Central Bank (ECB), which operated between 2011 and 2014, providing financial assistance program and imposing austerity measures and structural reforms to stabilize the Portuguese economy. Oliveira (2021) explains that, because of the crisis, between 2011 and 2016, Portuguese migration balances assumed negative values, with an increase in outflows and a slowdown in the inflow of new immigrants.

Another point to consider when choosing the years of analysis coverage was that there is still, as mentioned, limited empirical evidence regarding wage discrepancies between immigrants and natives in Portugal after 2010, with emphasis on studies such as those by Carneiro *et al.* (2012) and Cabral & Duarte (2012, 2016), who study the Portuguese market using QP for years before 2008.

Therefore, to study the existence of a wage difference between immigrants and native-born individuals in an up-to-date manner, the decision was made to utilize the last ten years available in the database, from 2010 to 2020. The study will investigate the recession period (2011-2013), the Troika intervention (2011-2014), the two recovery periods (the 2014-2016 post-Troika, and the 2017-2019 expansionary cycle), and examine if the discrepancy arises from the productive characteristics of jobs and workers. However, it is worth mentioning that one of the variables used in the study, as will be explained in the econometric analysis chapter, was used with a one-year lag, meaning that all results refer to the period between 2011 and 2020.

In addition, to eliminate missing or inconsistent values and avoid bias errors, the database underwent data cleaning. Firstly, duplicate records and workers whose identification numbers contained fewer than six digits due to data reproduction or filling problems were removed. Then, the sample was restricted to individuals aged 18 to 66, corresponding to workers without the regular age to access an old-age pension. The category of education "not defined", the level of professional qualification registered as "ignored", the NUTS II associated with "foreigner", stateless workers, and all observations of variables intended for the study and not included in the database were also eliminated.

Similarly, to maintain consistency within the sample, the data was restricted to full-time workers who had completed at least 140 monthly service hours. According to Portuguese legislation (Article 54 of the Labor Code), in cases such as for parents of children with disabilities, the regular service in Portugal may be reduced by five hours of work per week, equivalent to 35 hours weekly and 140 hours per month (Portugal, 2009). Employees who work less than this value and perform full-time jobs were therefore considered measurement errors. Moreover, according to Article 275 of the Labor Code, there is a legal possibility of a twenty percent reduction in the minimum monthly remuneration earned by practitioners, apprentices, interns, and workers in a situation of certified training (Portugal, 2009). For this reason, employees whose base salary was less than 80 percent of the minimum wage were eliminated.

Furthermore, it was also necessary to undertake the correction of inconsistent variables. Specifically, discrepancies were observed in data recording concerning the worker's tenure, age, and gender. For example, there were instances where individuals exhibited changes in gender classification from one year to the next, only to revert to the initially registered category afterward.

Finally, nominal remuneration data was adjusted to real values using the Consumer Price Index (CPI) provided by Statistics Portugal (INE)[1].

The processed data set includes 8,870,100 observations, averaging 887,010 data per year, with 45.98% female and 54.02% male. Among these figures, 8.14% represent immigrants, with 41.21% women.

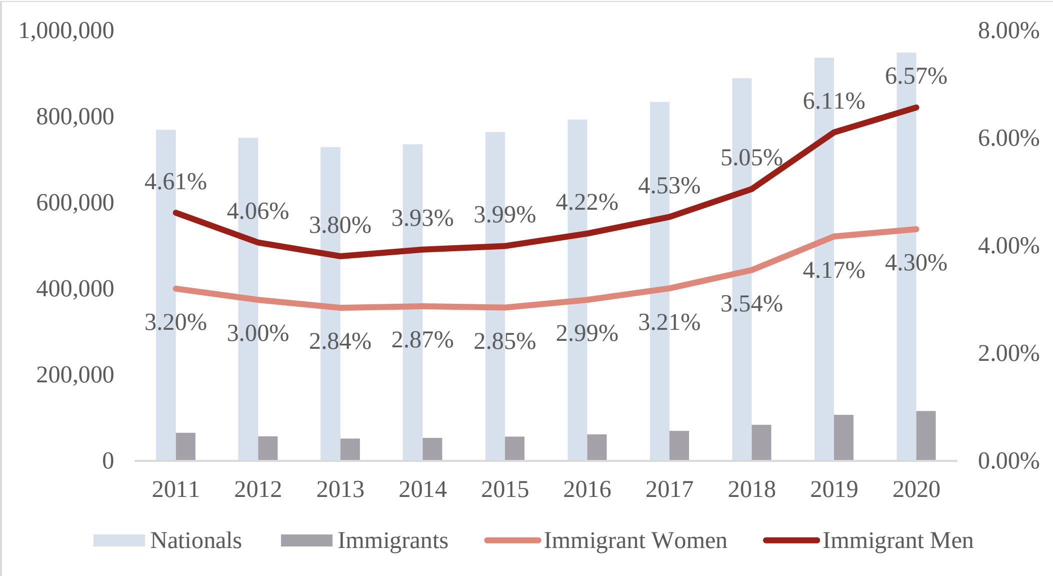
Figure 5 illustrates that the proportion of foreigners in the population of workers in the private sector in Portugal remains relatively small (721,941 immigrants compared to 8,148,159 nationals registered in the sample), with a more significant discrepancy in 2020. The first reduction in the number of foreigners can be seen in 2011, which reflects the consequences of the economic crisis in the country. The entry of non-national workers increased again, for both genders, from 2014, with the slowdown of the recession and the end of the Troika, with the most pronounced growth among men.

There was, however, a further slowdown in the number of immigrants between 2019 and 2020, particularly for females, with only a 7.83% increase in the sample of foreigners compared to the record figure of 28.30% in the previous year. Note that this reduction is not unrelated to the effects of the COVID-19 pandemic, which significantly restricted mobility between countries. In this context, it is worth noting that the QP reference period is October of each year. Therefore, the 2020 data already consider the beginning of the pandemic, a designation that began to be used by the World Health Organization (WHO) in March of that year.

For Portuguese workers, there was also a decrease in the number of observations after the onset of the 2010 crisis, with a reduction of 2.47% in 2012 compared to the previous year. Likewise, with the beginning of measures to contain COVID-19 in Portugal, there was also a drop in growth between 2019 and 2020, with only a 1.27% increase in native workers.

[1] Annex B shows Portugal's CPI evolution between 2011 and 2020.

Figure 5 - Number of foreign and native-born workers in Portugal (2011 - 2020)



Source: Quadros de Pessoal. Own elaboration.

3.3 Immigrants' origin

As mentioned earlier, Portugal has historically been a country of emigration, with a low relative importance of the foreign population compared to other European Union nations (Oliveira, 2021). However, this reality began to shift at the end of the 1990s, with an increase in demand for work by immigrants from Central and Eastern Europe (CEEC)[2], countries without privileged historical, cultural, or economic ties with Portugal (Cabral & Duarte, 2012). In addition, Iorio & Ferreira (2013) mention that, after the Portuguese recession period of 2010, with facilitating migration policies, micro-level issues (language and family strategies), and the worsening of Brazilian economic conditions in 2014, the increase of immigrants from the country was substantial, establishing itself as the largest foreign community in Portugal.

The Chinese community also stands out, with two distinct flows of immigration from China: the first, until the beginning of the nineties, related to Portugal's historical context with the former colony Macau, and the most recent, in the first decade of the century, characterized by economic immigration encouraged by laws that facilitate the legalization of foreigners (Gaspar, 2018). Finally, it is worth mentioning that the new wave of

---

[2] CEEC: Slovenia, Slovakia, Estonia, Hungary, Latvia, Lithuania, Moldova, Poland, Czech Republic, Romania, Russia, Serbia, and Ukraine.

immigration of workers from South Asian countries, including India and Nepal, has intensified since 2017.

Thus, in 2020, the main groups of foreign nationalities in Portugal, representing 90.03% of the total number of immigrants, were: Brazil, CEEC, China, Portuguese-speaking African countries (PALOP)[3], member countries of the European Union except Portugal (EU14)[4] and South Asia (SA)[5].

Supporting the earlier statement, Figure 6 shows the significant number of immigrants from Brazil, reaching 40,909 workers in 2020, a growth of 118.93% in nine years. However, South Asian immigrants are the most noteworthy, with an increase of 671.82% since 2011. Nepali-origin workers make up 40.65% of this group, with Indian workers following at 36.90%. It is important to note that the Portuguese government signed an agreement in 2021 to recruit Indian professionals, which is expected to increase even more the South Asian community in Portugal in the coming years (Portugal, 2021).

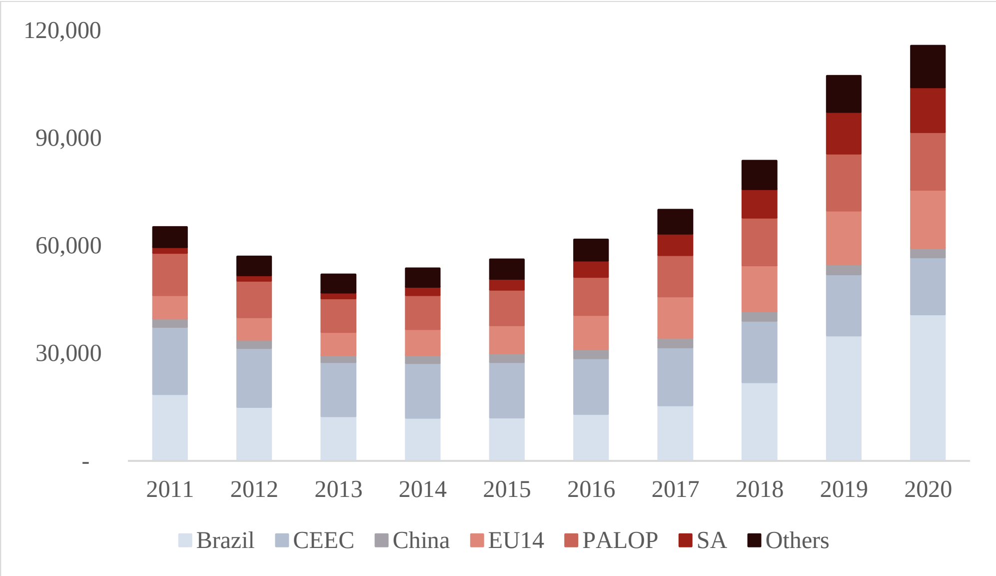
On the other hand, the EU14 and PALOP groups showed growth of 150.21% and 36.74%, respectively, between 2011 and 2020, revealing a significant increase in European workers and a more moderate growth in African residents. Oliveira (2021) explains that the slight increase in PALOP immigrants, compared to other groups, may be associated either with the acquisition of Portuguese nationality by these workers, since they are foreigners who have resided in the country for the longest time; or the departure of these citizens from Portugal due to economic conditions, and the support of programs such as "Support for Voluntary Return and Reintegration" benefiting migrants who wish to return to their country of origin and do not have the economic resources to do so. Finally, concerning these groups, it is worth noting that a significant proportion of EU14 workers originate from Spain (26.81%) and France (23.05%), while Cape Verde (41.94%) and Angola (26.12%) are the main origins for Africans.

[3] PALOP: Angola, Cape Verde, Guinea-Bissau, Mozambique, and São Tomé and Príncipe.

[4] EU14: Germany, Austria, Belgium, Denmark, Spain, Finland, France, Greece, Ireland, Italy, Luxembourg, Netherlands, United Kingdom, and Sweden.

[5] SA: Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, and Sri Lanka.

Figure 6 - Number of immigrant workers by group of nationality (2011 - 2020)



Notes: "Others" encompasses all foreign nationalities not included in other groups.

Source: Quadros de Pessoal. Own elaboration.

It's also important to note that Chinese workers had the most stable admission rate, with a minimum of 2,046 in 2013 and a maximum of 2,761 observations in 2019, the smallest group of workers in the sample. Finally, the percentage of CEEC workers was the only one that showed a reduction, reaching 15,840 observations in 2020, a negative difference of 15.12% since the beginning of the study period, ceasing to be the second largest group of immigrants in the sample and establishing itself in fourth position.

However, it is worth mentioning that the United Nations High Commissioner for Refugees reported that until February 2023, since the beginning of the invasion of Ukraine by Russia in 2022, Portugal had already granted 58,242 temporary protections to Ukrainian citizens and foreigners residing in Ukraine (UNHCR, 2023b). As 50.77% of immigrants from CEEC are of Ukrainian origin, it is expected that the number of workers in this group will increase significantly in the coming years, particularly women, due to compulsory male military service.

Lastly, Table 1 reveals that the proportion of men in the sample is higher than that of women for all groups, with South Asians having the highest male percentage of 92.07% in 2011. However, there was a constant increase in the proportion of women in employment for most analyzed groups, except for Brazil and PALOP, for which the values decreased between 2015 and 2020.

Table 1 - Gender distribution by year and nationality group

	2011		2015		2020	
	Women	**Men**	**Women**	**Men**	**Women**	**Men**
Portugal	44.68%	55.32%	46.63%	53.37%	46.86%	53.14%
Brazil	46.21%	53.79%	49.22%	50.78%	43.78%	56.22%
CEEC	36.19%	63.81%	38.20%	61.80%	41.41%	58.59%
China	39.78%	60.22%	39.84%	60.16%	42.05%	57.95%
EU14	43.89%	56.11%	45.07%	54.93%	46.25%	53.75%
PALOP	46.30%	53.70%	47.21%	52.79%	42.80%	57.20%
SA	7.93%	92.07%	10.69%	89.31%	12.54%	87.46%
Others	35.09%	64.91%	37.42%	62.58%	37.15%	62.85%

Source: Quadros de Pessoal. Own elaboration.

3.4 Endogenous variable

The real hourly wage was chosen as the endogenous term of the regression model to determine the difference in salaries between immigrants and nationals in Portugal. This variable, as done by Carneiro *et al.* (2012), is calculated by dividing the regular monthly remuneration corrected by the CPI by the total number of regular hours worked. However, in the labor economics literature, as in the Mincerian wage equation developed by Mincer (1974), the wage logarithm is often used as the response variable. This practice, as explained by Gujarati (2014), is adopted since wages do not usually have a normal distribution, which can affect the residuals, making them asymmetrically distributed.

It is important to note that the Ordinary Least Squares method, which will be implemented in this analysis, does not assume the probabilistic nature of the perturbation terms, which in turn may make it unfeasible to carry out statistical tests and inferences about the true coefficients of the population from the sample. To overcome this, a normal probability distribution is usually assumed because, in the presence of a large number of independent and identically distributed random variables (represented by the residuals), the Central Limit Theorem (CLT) of statistics shows that the sum of these terms tends to a Gaussian distribution (Gujarati & Porter, 2009).

Thus, the dependent variable of the study is computed as illustrated by Equation 1.

$$\ln Wageh = \ln\left(\frac{real\ base\ wage}{normal\ hours\ of\ work}\right) \qquad \text{(Eq. 1)}$$

Based on the data presented in Table 2, it is apparent that the average hourly wage of native workers surpasses that of immigrants. Specifically, Portuguese workers earn an average of 6.22 €/h, while foreigners receive 5.41 €/h. Immigrants from the European Union are an exception, as they have an average higher than that of nationals, corresponding to 10.50 €/h. Correspondingly, the residual group of nationalities ("Others") also exhibits a higher average than the Portuguese, 7.13 €/h. South Asians and Chinese, on the other hand, show the lowest wages, with averages of 3.58 €/h and 3.86 €/h, respectively.

Table 2 - Evolution of the real hourly wage by gender and group of nationality (in €/h)

	2011		2015		2020		Mean (2011 – 2020)		
	Women	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Δ%**
Portugal	5.60	6.81	5.60	6.65	6.00	6.88	5.67	6.69	-17.91%
Brazil	3.57	5.58	3.66	5.68	4.20	5.49	3.83	5.55	-44.83%
CEEC	3.45	3.92	3.62	3.95	4.73	4.87	3.86	4.25	-10.19%
China	3.23	3.24	3.41	3.62	4.66	5.01	3.71	3.96	-6.68%
EU14	7.77	16.28	7.54	15.71	7.12	12.16	7.20	13.30	-84.86%
PALOP	3.86	4.55	3.82	4.31	4.11	4.34	3.90	4.37	-12.09%
SA	3.52	3.36	3.26	3.26	3.95	3.86	3.65	3.57	2.07%
Others	4.64	8.04	4.69	8.50	5.35	8.08	4.89	8.50	-73.84%

Notes: the wage differential (Δ%) is defined as $\frac{\overline{W_w}-\overline{W_M}}{\overline{W_M}}$, where $\overline{W_w}$ is the real hourly wage for women and $\overline{W_M}$ for men.

Source: Quadros de Pessoal. Own elaboration.

In analyzing the data, it also has been noted that there exists a gender-based discrepancy in earnings. Except for the SA workers, women across all nationality groups earn less than their male counterparts, with the highest disparity observed amongst immigrants from EU14. The average salary gap in this group is as high as 84.86% compared to the male sample. However, it is noteworthy that the difference in remuneration between genders has decreased over the years. In particular, the European Union has shown improvement in this regard, with the gender gap reducing from -109.57% in 2011 to -70.88% in 2020. The only exception to this trend is the Chinese worker group, where the difference in earnings has increased by 7.05 percentage points over the past nine years.

Another important aspect is the percentage of workers earning minimum wage[6]. Table 3 shows that the Chinese group has the highest percentage of individuals receiving minimum remuneration, with over 60% of the sample falling under this classification. In contrast, Portuguese and immigrants from the European Union have the lowest average percentages. However, the proportion of natives earning minimum wage is higher than that of EU14 foreigners, underscoring the higher average salaries earned by workers from the European Union. Lastly, concerning gender, women are the majority group, especially among CEEC immigrants, in which 31.57% of the female sample earns the minimum wage. China and SA are exceptions, being the only groups whose male gender is more representative.

Table 3 - Evolution of minimum wage earners by gender and group of nationality

	2011		2015		2020	
	Women	**Men**	**Women**	**Men**	**Women**	**Men**
Portugal	9.29%	4.45%	10.74%	6.07%	16.29%	9.77%
Brazil	34.77%	23.15%	35.83%	26.65%	45.16%	34.09%
CEEC	33.98%	18.11%	36.16%	24.99%	34.15%	27.66%
China	64.98%	63.55%	66.32%	69.29%	64.90%	72.22%
EU14	8.90%	5.48%	8.63%	7.91%	7.74%	5.37%
PALOP	25.46%	16.12%	28.53%	23.18%	38.26%	32.59%
SA	41.41%	47.81%	50.15%	51.11%	50.86%	50.22%
Others	22.31%	18.01%	24.81%	21.67%	31.19%	29.93%

Source: Quadros de Pessoal. Own elaboration.

3.5 Exogenous variables

Regarding the factors influencing the real hourly wage, Mincer (1974) formulated a model wherein the natural logarithm of remuneration is expressed as a function of years of education, experience - generally estimated based on the worker's age - and a vector of observable characteristics of the individual, such as region and gender (Mincer, 1974; Lemieux, 2006). Given that this equation forms the foundation of labor economics, and in light of studies such as those conducted by Borjas (1994) and Carneiro *et al.* (2012), it was decided in this analysis to incorporate age, tenure in the firm, concentration of immigrants, level of education and qualification, location, economic activity and number of people working for the company as explanatory variables.

6 Annex B shows Portugal's minimum wage evolution between 2011 and 2020.

Furthermore, pioneering studies on immigration, such as those by Chiswick (1978) and Carliner (1980), include the number of years since migration (YSM) to help explain workers' wages and understand the process of assimilation in the host country. In the case of QP, however, it is not feasible to directly obtain the length of stay of foreigners in Portugal. Nonetheless, it is possible to calculate the year the worker legally entered the private sector and use this value as a proxy for the number of years since migration.

For this, two variables were calculated: for the first one, through the identification number of each worker, which is unique and constant over time, the year the individual was first registered in the database was sought. For the second, the length of time the worker had been with their company (tenure) was subtracted from the reference year of each observation.

Since individuals may have been hired before the beginning of the QP compilation, the minimum between these two variables, the year of first registration and the year of admission to a company, was used as the date of entry into the labor market. The year since migration was then calculated by subtracting the corresponding year from the minimum previously found. For native workers, this variable is null.

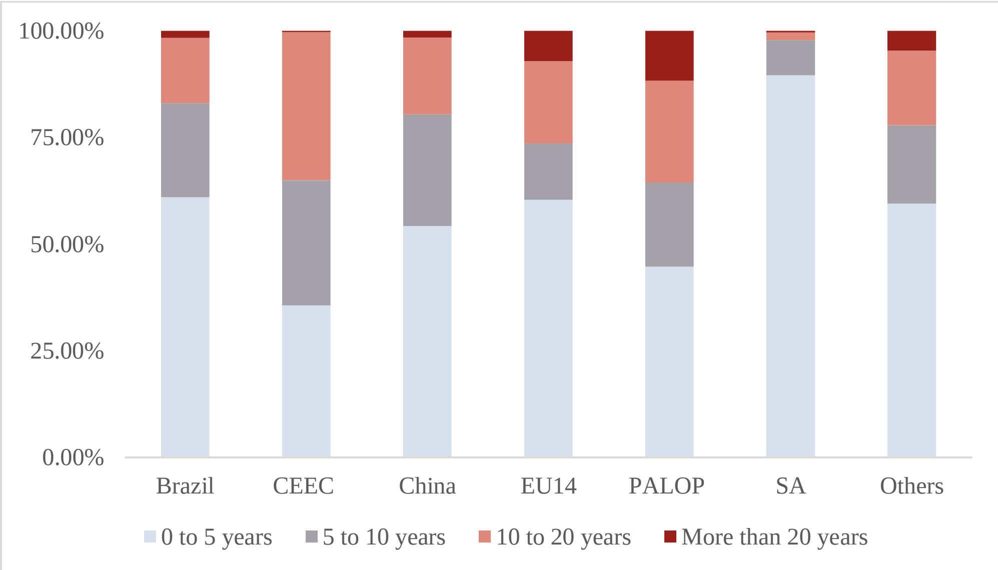
Although this approximation may not correspond exactly to the year immigrants entered Portugal, it is a valid method for minimizing the impact that the absence of this information may have on the results. This approach has been used in previous studies, such as those by Cabral & Duarte (2012) and Carneiro et al. (2012).

A detailed description of the variables for this study can be found in Annex A, while Annex B contains information on descriptive statistics, including average values by gender and nationality group.

Concerning the duration of residence following immigration, it merits attention that 54.25% of the foreign individuals within the sample integrated into the private job sector five years or less before the moment of their employment observation. Remarkably, 89.54% of South Asians account for these observed cases, corroborating the recent influx of this workforce into the country. Also, it is noteworthy that among all immigrant groups in Portugal, only South Asian women earn, on average, more than their male counterparts. This could be attributed to the group's recent arrival in the country. It is possible that during this initial period, women had better chances of securing jobs in sectors or occupations that are better remunerated in the Portuguese labor market, while men may have encountered more challenges in finding higher-paying employment opportunities.

Figure 7 also illustrates that PALOP presents the most substantial percentage of workers who have resided in the country for over twenty years, thereby representing immigrants with the lengthiest stay in Portugal. Conversely, CEEC displays the lowest proportion of workers with more than twenty years of experience in the country, comprising merely 0.28% of the observations within the group.

Figure 7 - Number of years since migration by group of nationality (2011 - 2020)



Source: Quadros de Pessoal. Own elaboration.

When analyzing the age groups of workers in Portugal, it is noticeable that the foreign population tends to be younger than the native. The average age for Portuguese workers is 39.99 years, while immigrants average 37.51 years. As an exception, CEEC male workers are older than the average of the Portuguese, with 40.32 years old, and are, therefore, the oldest group in the sample. South Asians, on the other hand, represent the youngest group, averaging 32.48 years old. It is also worth mentioning that, except for Brazil, women are usually younger than men, with Central and Eastern European immigrants showing the largest gender age difference of 2.10 years.

Additionally, Table 4 highlights a continuous increase in the average age of CEEC and Chinese workers from 2011 to 2020, indicating their permanence in the Portuguese labor market. In contrast, Brazil and EU14 showed an increase in the average age from 2011 to 2015, followed by a reduction, suggesting the entry of younger workers after the recession.

For native-born workers, the increase in the average age was consistently growing, indicating the aging of the population. This result is consistent with a study by Eurostat, which highlights the accelerated aging of the Portuguese population, with the third-highest median age among the 27 Member States of the European Union in 2020 (Eurostat, 2023d). This trend suggests that the national average age will continue to increase in the coming years. Despite this, immigration can act as a decelerator, with the entry of younger workers and the birth of babies to foreign mothers. However, it is important to note that, as mentioned, the influx of immigration in Portugal is still not as significant as in other EU states. Therefore, the impact of the entry of foreigners on the demographic structure of the country may not be as intense.

Table 4 - Evolution of workers' age by gender and group of nationality (in years)

	2011		2015		2020	
	Women	**Men**	**Women**	**Men**	**Women**	**Men**
Portugal	38.32	39.50	39.81	40.43	40.58	40.68
Brazil	34.90	34.51	37.69	36.79	36.55	35.32
CEEC	37.23	39.42	39.27	41.25	40.16	42.31
China	34.35	34.98	36.12	36.62	38.10	38.39
EU14	37.33	39.50	37.99	39.79	36.18	37.41
PALOP	36.96	38.53	38.45	38.41	37.33	37.37
SA	35.45	34.32	32.23	33.41	32.54	31.98
Others	37.60	30.09	38.84	38.79	37.51	37.11

Source: Quadros de Pessoal. Own elaboration.

Regarding employee tenure, defined as the length of time an individual has worked for an employer, national workers tend to remain with organizations longer than immigrants. Among foreigners, PALOP presents the highest values for tenure, with its workers maintaining an average of 4.91 years. South Asians, on the other hand, due to various factors, including their relatively short time in Portugal, remain approximately 1.04 years in firms. Additionally, except for Portuguese and workers from Central and Eastern Europe, women tend to have higher average tenures than their male counterparts.

Examining the evolution of values between 2011 and 2020, Table 5 indicates a downward trend in tenure averages following the economic recession for most groups. However, there was a continuous growth in tenure for workers from CEEC and China, which is consistent with the fact that these groups had the fewest new workers added to the sample.

Table 5 - Evolution of workers' tenure by gender and group of nationality (in years)

	2011		2015		2020	
	Women	Men	Women	Men	Women	Men
Portugal	8.69	9.39	9.68	9.96	9.22	8.89
Brazil	2.52	2.81	3.89	4.18	1.95	1.65
CEEC	3.26	3.79	4.70	4.84	4.80	5.25
China	2.21	1.86	2.64	2.30	3.35	2.98
EU14	5.28	5.04	5.30	5.05	3.55	3.47
PALOP	4.80	5.07	6.07	5.38	4.15	3.52
SA	3.00	1.68	1.60	1.23	1.12	0.94
Others	3.91	3.64	4.38	4.28	2.76	2.62

Source: Quadros de Pessoal. Own elaboration.

For immigrants' concentration in companies, this variable measures the proportion of non-national workers compared to the total number of people employed in an organization. It's a relevant factor in assessing the impact of foreign employees on both immigrant and national workers' wages.

On average, Chinese workers are employed in organizations with the highest concentration of immigrants amongst the sample. They are the only group working for companies with more than 50% of their workforce corresponding to non-nationals. This trend may be related to the specific economic activities of these immigrants, who usually work in wholesale and retail trades and accommodation and food service, which are sectors that historically employ many foreigners.

Table 6 - Evolution of immigrant concentration in firms by gender and group of nationality

	2011		2015		2020	
	Women	Men	Women	Men	Women	Men
Portugal	3.72%	3.80%	3.22%	3.39%	5.01%	5.47%
Brazil	27.02%	27.26%	23.73%	25.45%	23.60%	25.58%
CEEC	24.35%	25.07%	22.64%	25.58%	26.09%	28.69%
China	67.59%	68.75%	61.38%	63.59%	60.19%	63.86%
EU14	17.34%	17.11%	17.94%	17.87%	29.45%	28.77%
PALOP	15.76%	17.95%	15.30%	16.83%	17.58%	21.33%
SA	38.06%	48.46%	39.33%	46.93%	43.41%	45.68%
Others	21.77%	23.71%	22.64%	23.44%	26.30%	27.92%

Source: Quadros de Pessoal. Own elaboration.

It's also worth noting that Portuguese workers, on average, are employed in companies with a low concentration of immigrants, with roughly 97% of native employees. Finally, it's important to highlight that after the financial crisis, the proportion of foreigners in the workplace increased for most nationalities, except for Brazilian and Chinese women and South Asian men, for which there were minor decreases (see Table 6).

Regarding educational qualifications, all groups, including Portuguese nationals, exhibit primary education as the most prevalent category, except for workers from the European Union. As for EU14 immigrants the concentration lies in secondary education. This group also holds the highest proportion of individuals with postsecondary qualifications. Conversely, Chinese and South Asian workers demonstrate the lowest proportion of individuals holding a degree. Similarly, PALOP reports the lowest rate of masters and doctorates, with only 0.70% of female and 0.49% of male observations.

Analyzing the evolution of values, Table 7 demonstrates that the workforce in the Portuguese market is becoming increasingly educationally qualified. Specifically, the proportion of Portuguese natives and immigrant workers holding a degree increased by 4.09 and 6.37 percentage points, respectively, between 2011 and 2020. As a result, the number of workers with primary education as their highest completed literary qualification decreased by 13.45 and 17.68 percentage points among nationals and foreigners, respectively. Finally, it is worth noting that women dominate the college, master, and doctoral categories compared to men.

Table 7 - Evolution of foreign and native-born workers' education levels

	2011		2015		2020	
	Natives	**Immigrants**	**Natives**	**Immigrants**	**Natives**	**Immigrants**
Primary	51.72%	62.84%	44.94%	56.76%	38.26%	45.16%
Secondary	26.83%	26.33%	29.43%	29.41%	32.64%	35.68%
College	20.14%	9.77%	23.17%	12.21%	24.23%	16.14%
Mater	1.10%	0.73%	2.24%	1.29%	4.32%	2.28%
Doctorate	0.22%	0.33%	0.21%	0.33%	0.55%	0.73%

Source: Quadros de Pessoal. Own elaboration.

Concerning wages and educational qualifications, it has been observed that women with a doctoral degree receive the highest hourly wages of the female sample. This is also true for men, except China and PALOP, for which master's degree holders receive higher remunerations than workers from other categories. Likewise, as previously verified, EU14 immigrants are, on average, better paid than Portuguese workers, regardless of their educational background. However, it is interesting to note that the highest earning in the sample is not, in this case, for immigrants from the European Union, but for workers with a master's degree from China, given that Chinese men earn an average of 29.68 €/h. On the other hand, the lowest salary in the male sample also comes from China (3.34 €/h for workers with primary education). In case of women, CEEC immigrants have the lowest pay (3.33 €/h for qualifications below secondary education).

In occupational terms, the employment structure in the private sector reveals a greater concentration, for both genders and regardless of nationality, in the category of qualified and highly qualified professionals. Thus, 54.57% of native-born workers are concentrated in the third level of qualification, while for immigrants, the percentage is 63.41%. About the least significant category of the sample, the results reveal smaller percentages for the fifth level of qualification, that is, for apprentices, with 1.97% of native workers and 4.49% of foreigners.

In the sample, the most significant proportions of workers in management positions are found among Portuguese and European Union immigrants. Thus, for EU14, 48.16% of the male workers and 42.00% of the females are in supervisory and top executive roles. Considering that the citizens of this group come from countries whose minimum wages are, on average, higher than the Portuguese – in Spain and France, for example, countries that originate 49.86% of immigrants from the European Union, the minimum wage was, respectively, in 2020, 33.16% and 51.88% higher than that of Portugal -, and the fact that a high percentage of these individuals hold university degree, it is consistent that workers of this group accept positions with higher average wages (Eurostat, 2023a). Conversely, Africans and South Asians have the highest concentrations of workers in semi-skilled and unskilled categories (fourth level of qualification).

The Chinese, who have the highest percentage of minimum wage earners and a significant concentration of workers with primary education, show an interesting trend where most individuals occupy qualified and highly qualified positions, with 68.50% of their observations.

Finally, it is observed in Table 8 that during the economic crisis, the proportion of nationals in qualified, semi-qualified, and apprentice positions decreased. Similarly, there was a reduction in the number of immigrants with the two lowest qualifications, indicating that workers in higher management positions were better to maintain their jobs during the financial difficulties faced by Portugal.

Table 8 - Evolution of foreign and native-born workers' qualification levels

	2011		2015		2020	
	Natives	**Immigrants**	**Natives**	**Immigrants**	**Natives**	**Immigrants**
$Qual_1$	16.51%	7.56%	17.94%	9.01%	19.66%	8.99%
$Qual_2$	16.53%	6.69%	16.78%	7.61%	16.74%	7.64%
$Qual_3$	55.82%	62.16%	55.15%	64.60%	52.68%	60.64%
$Qual_4$	8.62%	17.24%	8.24%	14.65%	9.22%	18.90%
$Qual_5$	2.53%	6.35%	1.88%	4.14%	1.69%	3.84%

Source: Quadros de Pessoal. Own elaboration.

In regards to wages, as observed to previously variables, national workers do not receive, on average, higher wages than immigrants for all levels of qualification. Thus, for foreigners from the EU14, values remain the highest in the sample. South Asian women also earn more than nationals in semi and unskilled positions and apprentices. Additionally, Brazilian men receive salaries that are, on average, 125.87% higher than those of Portuguese in the second level of qualification.

Furthermore, for the female sample, it has been noted that salaries decrease as they approach the category of apprentices, with women in top executive positions, therefore, the best paid. Similarly, for men, the lowest wages are also recorded for the fifth level of qualification. The highest remunerations, on the other hand, vary according to nationality. Thus, for national, Chinese, African, and South Asian men, top executives have the highest salaries in the sample groups. As for Brazilians, Eastern and Central Europeans, and immigrants from the EU14, the second qualification is the best paid.

Regarding the salaries of apprentices, it should be noted that in Portugal, there are mandatory internships without remuneration for graduates, which generates losses for young professionals, unjustifiably postpones their entry into the labor market, significantly increases the costs of their training, and generates the lowest wages among the qualification levels. However, in 2023, adjustments were made to the Portuguese Professional Statutes, requiring a salary equivalent to the minimum wage plus 25% for this category. In practice, it will correspond to a minimum ceiling of € 950 per month (Portugal, 2023). Thus, it is expected that the average values recorded for this level of qualification will increase in the coming years.

Next, considering the geographical distribution of the population in the Portuguese territory and the job offers in Lisbon and the North, a greater concentration of national workers in both regions is expected, as confirmed by the results in Annex B. Thus, 75.94%

of Portuguese nationals work for companies in these geographic locations. As for immigrants, Lisbon is the main destination, concentrating 54.63% of workers. On the other hand, diverging from the nationals, the Center stands out with the second position, receiving 13.95% of the immigrant sample. Despite this, Table 9 shows that, between 2011 and 2020, the North was the location that received the most non-national workers, indicating growth of the region.

The Azores and Madeira, in turn, are the locations with the lowest concentrations of workers, receiving only 2.56% of Portuguese and 1.28% of foreigners. These two regions even showed a reduction in the number of individuals since the first year of the analysis.

When looking at the nationality groups, the female participants from CEEC had the highest percentage of workers in the Algarve region, with 21.85% of the observations. Additionally, the most representative immigrants in the Centre were from East and Central Europe, making up 25.43% and 19.02% of the male and female samples of the group, respectively.

Table 9 - Evolution of foreign and native-born workers' distribution across NUTS II regions

	2011		2015		2020	
	Natives	**Immigrants**	**Natives**	**Immigrants**	**Natives**	**Immigrants**
North	23.54%	11.80%	23.52%	12.47%	26.20%	14.21%
Algarve	3.30%	11.16%	3.31%	10.90%	3.35%	8.95%
Centre	14.27%	14.99%	13.96%	14.42%	15.72%	14.13%
Lisbon	52.35%	54.59%	52.79%	54.33%	48.55%	54.82%
Alentejo	3.60%	5.89%	3.86%	6.61%	3.76%	6.86%
Azores	1.04%	0.45%	1.03%	0.37%	0.97%	0.30%
Madeira	1.91%	1.12%	1.52%	0.91%	1.44%	0.72%

Source: Quadros de Pessoal. Own elaboration.

Associating salaries with the companies' geographical locations, it is evident that immigrants are, on average, better remunerated in the North and Madeira, with particular emphasis on European Union workers in the northern part of the country. Conversely, the Algarve and Alentejo regions record the lowest average wages among the foreign sample, with China and South Asia displaying the lowest compensations. Concerning natives, it is worth highlighting the wages in the Lisbon metropolitan area, with an average of 6.93 €/h, and also in Madeira, with 5.72 €/h. However, for the central part of the country, the average wage is the lowest within the Portuguese sample.

Concerning the wage disparity, it was found that South Asians show the smallest gender gap, with women being even better remunerated than men in most regions, except for Madeira. For CEEC in the Azores and China in Alentejo, the wage discrepancy is also positive in favor of the female sample. On the other hand, Brazilians have the largest negative differentials, with men receiving 192.55% and 112.55% more than women in Madeira and the Azores, respectively.

Regarding economic activities, Table 10 shows the sectors with the highest proportions of workers in the sample. Thus, manufacturing, wholesale and retail trade, transportation and storage, accommodation and food service, and administrative and human health activities represent 73.03% of Portuguese observations and 70.63% of immigrants. Among the activities, the highlight for nationals is in the third CAE, with 23.35% of the data. However, considerable growth between 2011 and 2020 has been observed in the observations of the health sector for this group, which may indicate the demand for professionals with the onset of the COVID-19 pandemic. As for foreigners, the most significant proportion is in accommodation and food service, corresponding to 19.43% of workers.

Among nationality groups, particular emphasis is warranted on China's prominent presence in the wholesale and retail trade, encompassing 54.02% of Chinese men and 68.91% of Chinese women. Additionally, men from CEEC and PALOP distinguish themselves from other samples by having a high number of workers in the construction sector. Finally, it is noteworthy that no employee is registered in the activities of households as employers (CAE_{20}).

Table 10 - Evolution of the five primary economic activities for foreign and native-born workers

	2011		2015		2020	
	Natives	**Immigrants**	**Natives**	**Immigrants**	**Natives**	**Immigrants**
CAE_3	23.06%	13.80%	23.42%	13.91%	23.50%	12.37%
CAE_7	15.86%	15.63%	14.90%	14.83%	15.49%	11.73%
CAE_8	7.23%	5.65%	7.05%	5.86%	6.05%	3.56%
CAE_9	6.07%	19.35%	6.20%	19.46%	5.90%	15.55%
CAE_{14}	8.52%	9.63%	9.30%	10.47%	9.03%	17.15%
CAE_{17}	9.71%	5.91%	11.79%	6.69%	13.14%	6.43%

Source: Quadros de Pessoal. Own elaboration.

About compensation, the average salary of national workers tends to be higher than that of immigrants across most categories. Despite this, the highest salary in the sample remains for individuals from the European Union, with entertainment activities being the most prominent category for this group, with an average wage of 133.44 €/h for male workers. Conversely, China and SA are the nationalities that register the least paid, with Chinese men receiving 3.26 €/h in accommodation and food service; and South Asian women 3.20 €/h in the waste management sector. For Portuguese nationals, on the other hand, extraterritorial organizations and electricity activities provide the highest remunerations in the group sample, while agriculture, administrative, and food service activities are the smallest.

Concerning gender, women receive the highest wage returns working with electricity and gas activities, with PALOP immigrants having the highest value in the female sample, 22.94 €/h. This result highlights the wage disparities between the category with the best remuneration in the male sample and the best in the female sample. In this aspect, South Asian workers are those who exhibit, for wholesale and retail trade and other service activities, the most equal remuneration between genders, with the discrepancy being less than 1%. Conversely, immigrants from the European Union who perform entertainment activities experience a significant difference of 1,615.62% between men and women in their salary.

Finally, regarding firm size, the concentration of national workers occurs mainly in large organizations, 57.01% of the group's observations. The distribution is more homogeneous for immigrants, with approximately 25% of the data in all categories. However, 28.53% of immigrants are also employed in large companies, making it the main category for this group. It's also noteworthy that there's an 18.23 percentage points difference between the proportions of foreigners and Portuguese in micro-organizations, with only 2.65% of nationals in such establishments. Table 11 even reveals a decrease in the percentage of Portuguese in companies with less than ten employees between 2011 and 2020.

Table 11 - Evolution of foreign and native-born workers' distribution across firm sizes

	2011		2015		2020	
	Natives	**Immigrants**	**Natives**	**Immigrants**	**Natives**	**Immigrants**
Micro	3.17%	24.49%	2.44%	23.32%	2.86%	16.47%
Small	14.34%	31.21%	12.62%	29.33%	14.18%	25.99%
Medium	26.90%	21.89%	27.06%	23.02%	27.61%	21.74%
Large	55.59%	22.40%	57.87%	24.33%	55.35%	35.80%

Notes: Micro (less than 10 employees), small (10 to 49 employees), medium (50 to 249 employees), and large (more than 250 employees).

Source: Quadros de Pessoal. Own elaboration.

It is also evident from Annex B, that Brazilians, Central and Eastern Europeans, and South Asians concentrate mainly on small-sized firms. In contrast, immigrants from the European Union and PALOP predominantly appear in companies with more than 250 employees. Interestingly, Chinese workers diverge from other groups, displaying significant proportions of individuals in micro-enterprises, accounting for 67.93% of male observations and 72.49% of female observations.

Lastly, it is expected for salaries that medium and large companies offer better compensation to their employees. Therefore, for nationals, wages are higher in larger firms, whereas for immigrants, the average is higher in medium-sized enterprises. Additionally, it is noteworthy that, once again, EU14 workers receive higher wages than Portuguese workers across all categories. Furthermore, it is interesting that China has the highest female salary in the sample (11.58 €/h in medium-sized companies) and the second-highest male salary (20.23 €/h also in medium-sized firms). Finally, concerning gender disparity, South Asian workers are the only group where female remuneration is equal to or higher than male salary across all company sizes.

3.6 Conclusion

The descriptive analysis carried out in this chapter provided a detailed view of the Portuguese private sector between 2011 and 2020, allowing a more comprehensive understanding of the variables that will be used in the forthcoming econometric investigation. Thus, although the immigrant population in Portugal has been increasing annually (with a pause in growth during the 2010 economic crisis), the number of foreign workers in the country is still limited. Additionally, the country's immigrants are a heterogeneous group whose characteristics vary significantly based on their origin, gender, and socioeconomic profile.

It was also evident that wage differences observed between the nationality groups reflect not only the specific characteristics of the immigrants but also factors related to firms, such as the level of professional qualification, geographic locations, and economic activities. Therefore, the analysis highlighted the need to apply techniques that allow an understanding of whether the wage discrepancy stems from the individual characteristics of workers or from the combinations they form with the firms where they work.

Furthermore, upon noting that the average remuneration of national workers is lower than that of European Union immigrants and also lower than that of other foreign nationalities, depending on the analyzed variable, it is expected that the results of the econometric analysis section will not indicate the persistence of negative effects on immigrant wages for all terms of interest.

The gender wage gap was also evident, especially for different economic activities. Thus, performing a stratified analysis on the database and segmenting the observations into groups based on the worker's gender will allow examining how the independent variables affect the wage logarithm differently for men and women. This subgroup approach, however, is not restricted to the gender of workers since relevant economic variables, such as NUTS, also help to identify whether there are specific effects on wages that can be attributed to each group.

The concentration of immigrants in firms is also a relevant variable that needs to be considered in understanding the salaries of foreigners and nationals, as it reflects the degree of diversity and ethnic heterogeneity in the Portuguese private labor market. China, for example, is a country whose workers concentrate in firms with an immigrant majority. Thus, understanding the effect of concentration on this group helps to comprehend whether the presence of immigrants leads to lower wages for Chinese workers compared to individuals who perform similar functions in companies with a higher proportion of Portuguese. For this, multiplicative terms between nationalities and concentration are a relevant approach to the research.

From the study of other important variables, the descriptive analysis also highlighted that the most recent immigration flow to Portugal comes from South Asia, with the workers from this group being the youngest and having the lowest tenure average in companies. On the other hand, Central and Eastern Europeans are the oldest immigrants in the sample, while Africans reside in Portugal for the longest time. China, PALOP, and SA are the least educationally qualified groups, and the EU14 has the most university-educated workers. Furthermore, Lisbon is the main workplace for Portuguese natives and immigrants, with the North and Center standing out in second and third positions. Finally, the third level of professional qualification, large-sized firms, and the manufacturing and accommodation and food service sectors stand out with a substantial number of observations in the sample.

In conclusion, the descriptive analysis not only summarized the main patterns of the database but also identified the most suitable statistical methodologies for the investigation in question, which will be addressed in the next chapter.

4 Methodology and econometric analysis

Access to a database containing information on the relationships between workers and companies, along with data regarding the gender and nationality of individuals, will allow, as elucidated in the preceding chapter, a comprehensive assessment of the wage divergence between immigrants and nationals in the Portuguese private sector between 2011 and 2020, and the examination of sample differences by gender.

Furthermore, considering that the descriptive analysis indicated a sensitivity of the wage disparity dimension to the set of explanatory variables, this research chose to progressively add new covariates to the regression to understand the source of the divergence. However, applying a decomposition technique is necessary since the simple comparison between specifications does not identify population parameters of interest (Gelbach, 2016).

A supervised regularization methodology was subsequently applied to enhance prediction accuracy and the interpretability of the resulting statistical model. This technique enables the comparison of covariates and the selection of only the most relevant ones, thereby eliminating irrelevant information while minimizing the prediction error (Freijeiro-González *et al.*, 2022).

4.1 Wages in the Portuguese private sector

Based on the variables highlighted in the previous chapter, the analysis will commence by considering the first static panel model – Pooled OLS (POLS) – to evaluate the behavior of the dependent variable. Under POLS, explanatory variables are assumed to have a common impact across workers and firms. Additionally, following the approach undertaken by Carneiro *et al.* (2012), different specifications were formulated to ascertain whether the wage differential between native and immigrant workers with similar characteristics arises from the attributes of the companies or the interactions they form with individuals.

Thus, the first specification considers a set of regressors that only control workers' characteristics (YSM, age, and education levels), with the application being carried out for two distinct samples according to gender. It is important to emphasize that region dummies, corresponding to NUTS II, were included in all regressions to capture the average impact of working in a specific region on the salary.

$$\ln Wageh_{it} = \beta_0 + \beta_1 \cdot Immigrant_{it} + \beta_2 \cdot YSM_{it} + \beta_3 \cdot YSM_{it}^2 + \beta_4 \cdot Age_{it} + \beta_5 \cdot Age_{it}^2 + \sum_{j=2}^{5} \beta_{6j} \cdot Educ_{jit} + \beta_7 \cdot Algarve_{it} + \beta_8 \cdot Centre_{it} + \beta_9 \cdot Lisbon_{it} + \beta_{10} \cdot Alentejo_{it} + \beta_{11} \cdot Azores_{it} + \beta_{12} \cdot Madeira_{it} + \varepsilon_{it}{}^{7} \quad \text{(Spe. 1)}$$

In Specification 1, it is evident that a dummy variable representing immigrant status has been included to the equation in addition to the variables mentioned during the exploratory analysis. This addition enables the separate estimation of the nationality's effect on wages for Portuguese and foreign workers. Given the research objective, "Immigrant" becomes the variable of primary interest for the analysis. Furthermore, the squared terms of the age and "years since migration" regressors were also included, allowing for a non-linear relationship between salary and these variables, thus facilitating the capture of possible accelerations or decelerations in the effect of worker aging and duration of stay in Portugal on the dependent term. Lastly, it is worth noting that the reference categories for the binary variables are described in Annex A. For Specification 1, primary education and North are the omitted terms.

An important assumption to be considered before estimating the equations is the variability of the disturbance terms. In presence of heteroscedasticity, characterized by the non-constant dispersion of the residuals, the OLS estimators do not have minimum variance, nor are they efficient. This issue arises because the variance of the coefficients, under the assumption of homoscedasticity, as demonstrated by Dhrymes (1978), is usually smaller compared to the variance in the presence of heteroscedasticity, meaning that confidence intervals based on the latter will be unnecessarily wider (Gujarati & Porter, 2009). Consequently, the t-tests and F-tests lose their reliability, leading to erroneous conclusions concerning the statistical significance of the estimated coefficients.

For this reason, it has been decided to employ the Breusch-Pagan test, which conducts an auxiliary regression on the squared residuals to detect whether the variance of the disturbance terms is associated with the predictor variables of the model (Gujarati, 2014).

The test's outcome revealed a p-value for the chi-square statistic below 0.05 across all specifications (see Annex C), indicating that, at a significance level of 5%, the null

[7] β denotes the coefficients that quantify the relationship among the predictors and the endogenous variable; and ε_{it} represents the error term for observation i at time t.

hypothesis of constant variance is rejected. Therefore, to address this issue, the White method was utilized, which allows for the adjustment of coefficient standard errors in large-scale samples – as is the case of this study – restoring the validity of statistical inference (White, 1980; Gujarati, 2014).

To achieve this, White (1980) proposes the use of residuals ($\widehat{w}_j$) obtained from an auxiliary regression of the regressor X_j on the remaining exogenous variables; and the replacement of the variance of the disturbance terms - used in the homoscedastic formula of variance of the coefficients - by the squared residuals of the original regression ($\hat{u}_i^2$). In this way, Equation 2 of robust variance is obtained, which the author demonstrated to be a consistent estimator of the original formula under any form of heteroscedasticity.

$$\widehat{var}(\hat{\beta}_j) = \frac{\sum \widehat{w}_{ji}^2 \hat{u}_i^2}{(\sum \widehat{w}_{ji}^2)^2} \quad \text{(Eq. 2)}$$

Annex D presents the estimation results, wherein it can be observed that the "Immigrant" coefficient in Specification 1 is statistically significant at the 1% level and equal to -0.0731 in the case of the female sample and -0.0977 in the male. This indicates that the average hourly wage of a foreign worker, with primary education and employed in a firm located in the Northern region of Portugal, at the time of their arrival in the country, is approximately 7.05%[8] lower than that of Portuguese women, and 9.31% lower than that of native men with similar characteristics.

The signs of the coefficients for "years since migration" and its quadratic term reveal, however, that initially, as assimilation occurs, the average wage of immigrants decreases until reaching the inflection of the parabola. At this point, there is a reversal in the trend, and each additional year of stay leads to a less negative coefficient, reducing the wage penalty. Thus, for women, setting the first-order partial derivative of the logarithm of wages with respect to YSM[9] equal to zero yields a value of 10.50, indicating that after approximately ten and a half years of migration, each additional unit of time spent in the

[8] $\left(e^{\hat{\beta}} - 1\right) \cdot 100\%$ is the percentage change in real hourly wage associated with a one-unit increase in the independent variable. $\hat{\beta}$ is the estimator of a regression coefficient.

[9] $\frac{d \ln Wageh}{dYSM} = \beta_2 + (2 \cdot \beta_3) \cdot YSM$ is the first-order partial derivative of $\ln Wageh$ with respect to YSM, which, when equal to zero, returns the parabola vertex.

country reduces the negative effect on the wages of the female sample. For men, this value is equivalent to 13.15 years, suggesting that they enter the country with a more discrepant wage compared to nationals than women and take longer than them to experience a reduced wage penalty for the duration of their stay in Portugal.

Concerning age, the findings indicate a positive coefficient sign for both genders, reflecting an annual wage increment of approximately 3.70% and 4.38% for women and men, respectively. Furthermore, age demonstrates a negative quadratic effect, suggesting that, as workers age, the logarithm of wages initially experiences an upward trend, with this relationship becoming less positive as the predictor variable increases. This pattern suggests an age range where remunerations are maximized (in this case, 64.44 years for women and 66.97 years for men), beyond which wage growth decelerates. Therefore, it is noteworthy that the positive impact of age on male remuneration surpasses that observed for females, and the deceleration in wage growth occurs later for men.

When controlling for educational level, the higher the academic degree, the greater the positive effects estimated on workers' wages. Interestingly, the returns are more significant for men, except for the doctorate category, for which women present a higher coefficient. Thus, the salary of female doctorate holders is approximately 253.18% higher than that of women with primary education. For men, the difference between these two categories is 238.31%.

Finally, with the highest coefficient among the NUTS binary variables, Lisbon compensates its female workers 9.53% better than the Northern region (8.36% for the male sample). For the other NUTS, except for Algarve for men, Alentejo for women, and Centro for both, wages are also, on average, higher than the omitted category, considering all other factors constant. This result is noteworthy, given that the Northern region is an important industrial and service center with the second-largest population concentration in the country.

It is noted that the first specification does not consider time-fixed effects. However, as mentioned earlier, the study also aims to investigate the impact of economic recession and expansion periods on wages in the Portuguese market. Therefore, considering that the cost function changes over time due to influences like financial crises, shifts in government regulations, immigration policies, and comparable effects, binary variables representing the study's coverage years were also incorporated into the regression analysis, resulting in the formulation of Specification 2, and the application of the method POLS with time-fixed effects. It is of note that the base category utilized for the years dummies was 2011.

$$\ln Wageh_{it} = \beta X_{it} + \delta_t + \varepsilon_{it}{}^{10} \qquad \text{(Spe. 2)}$$

From the results of the second specification, it is observed that the signs of the coefficients remained unchanged compared to Specification 1, and the magnitude of the parameters underwent minor alterations for both gender samples. This means that the annual policies implemented by the government of Portugal, whether aimed at attracting immigrants or related to the country's economy, have a negligible impact on the quantitative variation concerning the effect of each explanatory variable on wages. Although the results suggest that the consideration of time-fixed effects may be deemed unnecessary, the decision was made to retain them in the model specification. This choice was motivated by the aim to enable a direct comparison of the findings of this study with those encountered in the Portuguese literature, such as the study by Carneiro *et al.* (2012).

In this manner, continuing the interpretation of the results, it is observed in Specification 2 that, in 2011, the wage of an immigrant with primary education located in the Northern region of the country was 7.66% and 9.34% lower than that of native women and men, respectively. Furthermore, for the temporal coefficients, wages are negatively affected for the years represented by the binary variables compared to the base year. Thus, the female logarithm of wages was approximately 5.79% lower in 2015 than in 2011. From 2015 onwards, the wage penalty decreased, reaching -0.48% in 2020. In contrast, the male wage penalty reached its maximum in 2017 at -8.38% and then decreased to -4.05% in 2020.

These values are relevant as they indicate that, until 2015 (2017 for the male sample), the negative effect on wages for workers in the Portuguese private sector, while considering all other factors constant, continued to increase relative to the beginning of the financial crisis. However, even after 2015 (2017), wages did not exhibit positive effects, reflecting the prolonged impact of the economic crisis in the country.

[10] X_{it} is a vector of predictors from Specification 1 (for entity i at time t); while δ_t is the unknow coefficient for the time regressors (t).

Also, as mentioned, it is important to compare the results obtained in this specification with the findings in the existing literature. Many studies, such as Oaxaca (1973), Beach & Worswick (1993), and Shamsuddin (1998), consider family-related characteristics, such as marital status and number of children, to assess workers' wages, especially when investigating gender disparities. Although QP does not allow control of such variables, Specification 2 is considered to be similar to those found in many analyses.

Schoeni (1997), for example, found that Mexicans aged 25 to 34 years, when entering the United States in the 1960s, earned about 38% less than native male workers, with this figure increasing to -43% in 1990, and representing a five percentage point decline in wages over twenty years. Friedberg (2000), on the other hand, in a study conducted for the Israeli market with workers aged 25 to 65 years (although not distinguishing by gender), found a divergence of -22.35% in the wages of foreign workers upon arrival, with their remuneration increasing by 0.80% per year of assimilation. Lastly, in a more recent analysis for the period 2003-2008, Carneiro *et al.* (2012) concluded that, upon entering Portugal, immigrants aged between 18 and 60 years, with less than six years of formal education, received approximately 24.90% and 16.30% less than native men and women, respectively.

Thus, when comparing the study of Carneiro *et al.* (2012) with the findings obtained in this analysis, considering the similarity in the specifications, there is a significant reduction in the salary difference between immigrants and natives in Portugal reported by both works. Also noteworthy is the divergence in the signs obtained for the variable YSM and its quadratic term. According to Carneiro *et al.* (2012), immigrants experience wage growth in the early years post-arrival, but the increase decelerates after reaching its peak.

Subsequently, by adding a vector of characteristics related to companies to regression 2, such as immigrant concentration, the number of employees, and economic activity, Specification 3 is obtained. It is important to mention that, to minimize potential endogeneity issues, the variable concentration of immigrants, as done by Carneiro *et al.* (2012), was used with a one-year lag. Thus, for company k in year t, the concentration variable measures the proportion of foreign workers in year $t-1$. Also, the base category for economic activity is the third CAE, manufacturing.

$$\ln Wageh_{it} = \beta X_{it} + \delta_t + \gamma c_{it} + \varepsilon_{it}{}^{11} \qquad \text{(Spe. 3)}$$

The addition of control variables for firm characteristics reduced the estimated "Immigrant" coefficient to -0.00933, resulting in a wage penalty of -0.93% upon arrival for female foreigners with primary education in the Northern manufacturing sector. As for men, it is interesting to note that the dummy variable representing foreign status is no longer significant, indicating that the company-related variables capture part of the variation previously attributed to "Immigrant" for this gender.

In turn, the "year since migration" variable maintained the signs from the second specification, showing increased inflection points. Thus, the female wage begins to exhibit less negative effects after 11.23 years of migration, while for men, this point is established at 13.94 years.

Moreover, the coefficients of age and its quadratic term showed consistent signs as obtained previously, but different trends to the apex of the parabola according to gender. For the female sample, the maximum point increased to 66.16 years, whereas for the male, it decreased to 65.52 years. Consequently, the women's wage deceleration was postponed by 1.95 years, while the males' salary was anticipated by 1.83 years.

Regarding educational qualifications, the coefficients remained positive, and the most substantial wage increments were still observed for workers with a doctorate. For the year dummies, the result of wage penalties for records in years other than the base year, 2011, remained unchanged, except for the female sample in 2020. Thus, the wage of female workers in the final year of the sample is 1.92% higher than that of women registered in 2011.

As for firms' locations, the variables remained positive for most regions, except for the Azores in the male sample and the Center for both genders, confirming again the negative effect on workers' wages in the Northern part of Portugal.

Concerning the variable concentration of immigrants, negative coefficients were observed for both genders, revealing that a higher proportion of immigrants in companies in the previous year leads to more pronounced negative wage effects in the current year. Specifically, a 1% increase in foreign workers results in reductions of approximately 2.03%

---

[11] c_{it} is a vector including the variables concentration, size, and the CAEs; while γ is a vector of coefficients that quantify the relationship between the variables in c_{it} and the wage.

and 3.03% in the wages of the female and male samples, respectively. Conversely, the size of organizations exhibits a positive impact, as expected, with workers in larger companies receiving superior remuneration.

Finally, in this specification, the coefficients related to economic activities are worth highlighting. For the male sample, they are mainly negative, revealing that workers in sectors like agriculture, administration, and human health receive lower salaries compared to those in the manufacturing industry. On the other hand, women show a higher proportion of positive coefficients, indicating that wage returns in many sectors, such as construction and education, surpass those of the omitted class. Additionally, certain dummy variables exhibit divergent coefficients between genders. Consequently, women experience positive wage rewards in construction and transportation, while men face negative wage effects.

Lastly, with the addition of variables related to the combinations formed by companies and individuals (match characteristics), such as tenure and qualification level, Specification 4 is formulated. Analogous to the approach taken with age and years since migration, the variable of tenure squared was included, thereby accommodating a non-strictly linear relationship between wages and the duration of a worker's employment within an organization. Also, for the qualification level dummies, apprentices is the omitted category.

$$\ln Wageh_{it} = \beta X_{it} + \delta_t + \gamma c_{it} + \theta m_{it} + \varepsilon_{it} \quad \text{(Spe. 4)}$$ [12]

An interesting result is observed for the coefficient of the "Immigrant" variable in this regression. For both genders, the results are positive and statistically significant, indicating that upon arrival, immigrant workers receive, on average, higher wages than nationals with similar characteristics (considering the base categories: 2011, primary education, Northern region, manufacturing, and apprentices). The percentage is 4.19% for the female sample, and 4.78% for the male. This finding diverges from that obtained by Carneiro *et al.* (2012), whose study, in a similar regression, identified a negative differential of 4.30% for Portuguese women and 7.92% for native men.

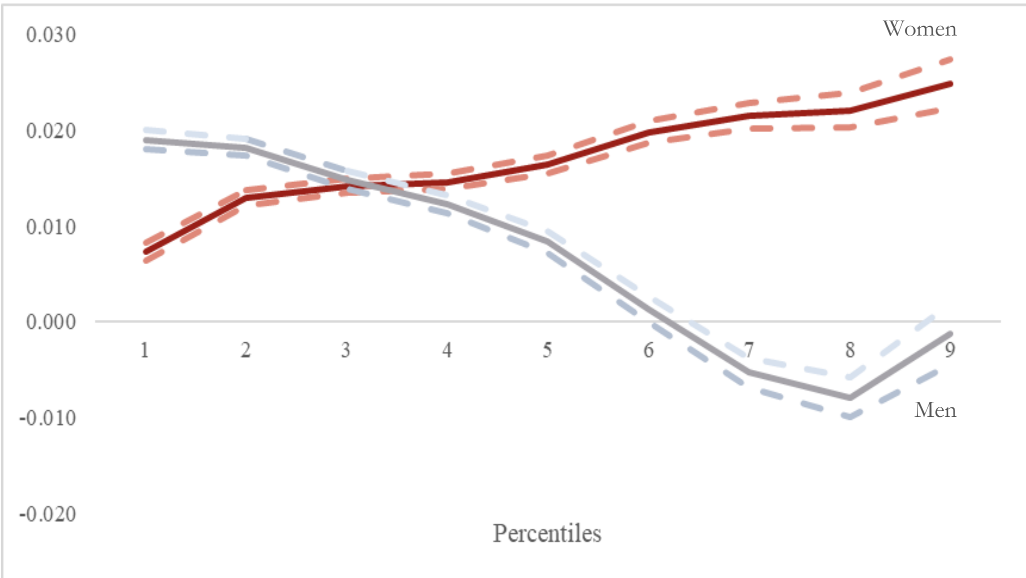
Thus, the divergence between the results obtained for "Immigrant" in regressions 2 and 4 indicates that factors related to match characteristics can positively influence the wages of

[12] m_{it} is a vector including the variables tenure, tenure squared, and the qualification levels; while θ is a vector of coefficients that quantify the relationship between the variables in m_{it} and the wage.

foreigners. Nevertheless, it is recognized the importance of implementing a decomposition method that allows an understanding of the role of covariates in the variation of the coefficient of the base regressor, as the mere comparison between specifications does not account for sensitivity to the order in which variables are added. Consequently, the next section will implement Gelbach's decomposition (Carneiro *et al.*, 2012; Gelbach, 2016).

However, it is possible to ascertain, through Specification 4 quantile regressions, that the wage penalty for native male workers is not constant across all percentiles (see Figure 8). Thus, for this group, the wage penalty decreases between the first and eighth deciles and starts to increase in the last one. The estimated coefficient even becomes negative in the seventh decile, indicating penalized remuneration within the top 30% of the wage distribution for immigrant workers. For women, on the other hand, the penalty for native workers increases across the entire wage distribution, ranging from a minimum of 0.00733 to a maximum of 0.0249 in the ninth decile.

Figure 8 - Estimated coefficients for "Immigrant" obtained through Specification 4 quantile regressions



Notes: (i) Solid lines represent the estimated coefficient at different deciles.
(ii) Dotted lines represent a 95% confidence interval.
Source: Quadros de Pessoal. Own elaboration.

Regarding the years since migration, the observed signs from specifications 2 and 3 were maintained, with an increase in the inflection point to 15.85 and 16.39 years for women and men, respectively. Thus, for all regressions, the results confirm a negative initial assimilation effect on wages, followed by a reversal of the trend from the minimum of the parabola.

For age, the positive impact on salary decreased for both genders. On the other hand, the negative and statistically significant quadratic effect remained, reinforcing the non-linearity between the variables, and the deceleration of wage growth starting at 57.10 years for women and 57.99 years for men.

As for educational qualifications, doctorate remained the category with the highest rate of return compared to workers with primary education, regardless of gender. However, the male wage return in the college's degree category was found to be higher than that for the master's degree, although the difference was not very significant. Thus, male graduates receive 49.33% more than individuals with primary education, while masters have salaries 49.03% higher.

Regarding the years, unlike the third regression, the coefficient of the male sample for 2020 became positive, indicating that for both genders, workers' wages show a favorable increase in the last year of the analysis compared to 2011. However, by examining the estimates of the quantile regressions in Annex E, it is possible to ascertain that the coefficients between 2016 and 2019 are not consistently negative, showing positive signs for the first deciles of the wage distribution. In summary, starting from 2016, there was a progressive improvement in wages, with the remuneration of workers initially located in the lower percentiles being higher than those observed in 2011.

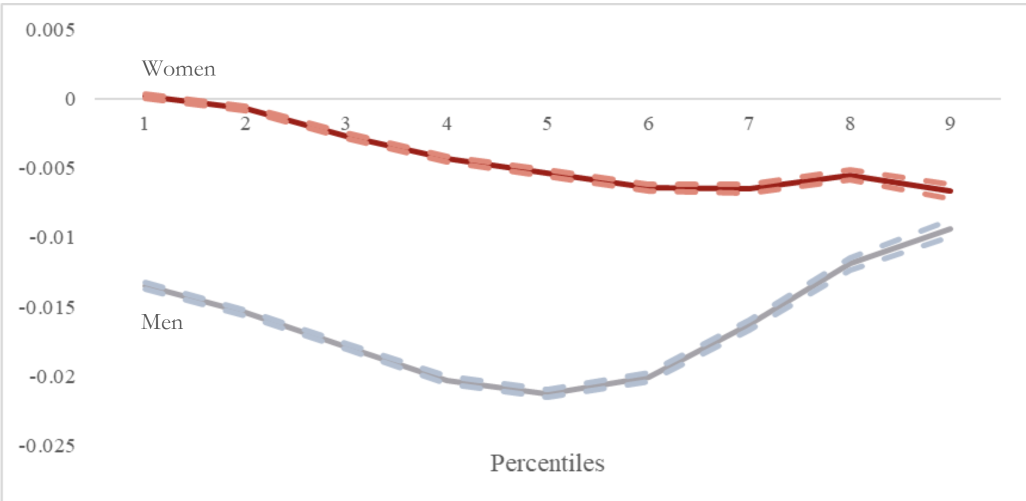
As for the NUTS, the sign pattern remained consistent compared to Specification 3, establishing Lisbon as the region where, on average, the best wage effects are observed compared to the North. The quantile analysis confirms positive coefficients for the geographic location dummy variables across the entire female wage distribution, with the exception of the Central region. However, for men, there are negative effects in the eighth decile for the Central region and the lower 80% of the wage distribution in the Azores.

Upon conducting the exploratory analysis, it was anticipated that Lisbon and Madeira would exhibit positive coefficients due to their higher average wages in certain worker groups compared to the North. However, the results of Alentejo and Algarve diverged from the initial analysis as they had the lowest average wages in the sample but still demonstrated positive coefficients. Nevertheless, it is important to note that the binary variable coefficients do not convey descriptive wage differences; rather, they signify the association between the regions and wages after controlling for other predictor variables.

Next, concerning the concentration of immigrants, the coefficients remained negative, indicating that a 1% increase in the proportion of foreign workers in firms reduces wages in the male sample by -0.0125% (-0.00811% for the female sample), therefore not having a very high impact on the remuneration. However, in Figure 9, it is notable that the magnitude of the concentration effect is not constant in the wage distribution, with a growing negative effect for women and an increasing effect for men up to the fifth decile.

The apparent contradiction between the positive coefficient of "Immigrant" and the negative coefficient of "concentration" can be explained by considering different levels of analysis. At the individual level, on average, foreign workers have higher wages than nationals with similar characteristics. However, at the company level, a high concentration of immigrants may indicate the presence of sectors with lower remuneration or less skilled job positions, where salaries are naturally lower. Moreover, as mentioned, some studies, such as those by Carneiro *et al.* (2012) and Andersson *et al.* (2014), suggest a tendency for cultural proximity when choosing a workplace. Thus, immigrants may be more willing to work in companies with other foreigners, especially if they share commonalities such as language and work habits, often accepting lower salaries under these circumstances.

Figure 9 - Estimated coefficients for "Concentration" obtained through Specification 4 quantile regressions



Notes: (i) Solid lines represent the estimated coefficient at different deciles.

(ii) Dotted lines represent a 95% confidence interval.

Source: Quadros de Pessoal. Own elaboration.

Additionally, the firm size variable remained positive for both genders; however, there was a reduction in the favorable effect on wages compared to other specifications.

Through quantile regressions, it is even possible to observe that from the seventh decile (sixth for men), the coefficient becomes negative for the female sample, indicating that for workers concentrated in the upper deciles, an increase in the number of employees leads to wage reductions.

As for economic activities, the coefficients for women in the construction and education sectors, and men in the information and consultancy areas, have changed direction, suggesting a negative impact on wages compared to the manufacturing industry. On the other hand, popular sectors, such as accommodation and food service and wholesale and retail trade, continued to have negative signs for both genders.

Regarding tenure, the coefficient was positive for both women and men, indicating that the longer a worker stays in a company, the higher the expected wage returns. However, the quadratic term showed a negative sign, suggesting that wages tend to increase as the years of employment in a company grow but at a decreasing rate. In other words, workers generally experience significant remuneration increases in the early years, and as they accumulate more years of work, subsequent increments become smaller.

Lastly, the highest positive effects of qualification level on wages occur for top executives. On the other hand, the wages of semi-skilled and unskilled workers are penalized compared to the base category. Positive effects are observed for the fourth qualification level only for the male sample, but this is limited to the bottom 40% of the salary distribution.

4.2 Gelbach's decomposition

Estimating different specifications for the same linear regression model, as done in this study, is a common practice in the literature. Gelbach (2016) explains that several authors, such as Yavuz *et al.* (2020), formulate base regressions and then progressively add new covariates to evaluate the change in the coefficient of interest. However, as mentioned earlier, the author questions this approach, emphasizing that the order in which the covariates are added can affect the results. Only if the variable of interest ("Immigrant" in this case) is orthogonal to the covariates, that the order will be irrelevant, as there will be no correlation between "Immigrant" and the other terms included in the regression model (Carneiro, 2021). Thus, the author proposes a decomposition method based on the omitted variable bias formula that allows to investigate the impact of each covariate to the coefficient change.

More specifically, the approach recovers the individual contribution of each additional variable, which is obtained conditionally on all other covariates. That is, Gelbach (2016) suggests first estimating the full regression (with all covariates of the analysis, including the variable of interest) by OLS, which will return the vector of coefficients $\hat{\beta}_k = (X'X)^{-1}X'Y$. Then, estimate auxiliary regressions for the covariates X_{2k}, where they become the dependent term, and the variable of interest X_1 becomes the predictor. The result is the vector of coefficients $\hat{\Gamma}_k = (X_1'X_1)^{-1}X_1'X_{2k}$, which, when multiplied by $\hat{\beta}_{2k}$, returns $\hat{\delta}_k$ (Equation 3), the part of the sample omitted variables bias attributable to covariate k when estimating $\hat{\beta}_1$.

$$\hat{\delta}_k = \hat{\Gamma}_k\hat{\beta}_{2k} = (X_1'X_1)^{-1}X_1'X_{2k}\hat{\beta}_{2k} \qquad \text{(Eq. 3)}$$

Because the sample omitted variables bias formula is linear in its k_2 components, it follows that the population omitted variables bias ($\hat{\delta}$) is the combined effects of individual biases, as shown by Equation 4.

$$\hat{\delta} = \sum_{k=1}^{k_2} \hat{\Gamma}_k\,\hat{\beta}_{2k} = \sum_{k=1}^{k_2} \hat{\delta}_k \qquad \text{(Eq. 4)}$$

Also, as Gelbach (2016) explains, the population omitted variables bias formula and the difference between the base (excluding all covariates of the analysis, leaving only the variable of interest) and full specification coefficient estimates are the same. As proof, the author demonstrates that multiplying both sides of the full specification by the projection matrix, $(X_1'X_1)^{-1}X_1'$, leads to the derivation of Equation 7.

$$\underbrace{(X_1'X_1)^{-1}X_1'Y}_{\hat{\beta}_1^{base}} = (X_1\hat{\beta}_1^{full} + X_{2k}\hat{\beta}_{2k} + \hat{u})(X_1'X_1)^{-1}X_1' \qquad \text{(Eq. 5)}$$

$$\hat{\beta}_1^{base} = \hat{\beta}_1^{full} + \underbrace{(X_1'X_1)^{-1}X_1'X_{2k}\hat{\beta}_{2k}}_{\hat{\delta}} \qquad \text{(Eq. 6)}$$

$$\hat{\delta} = \hat{\beta}_1^{base} - \hat{\beta}_1^{full} \qquad \text{(Eq. 7)}$$

For this analysis, Specification 2 corresponds to the base regression and Specification 4 to the full equation since, in this way, it is possible to investigate the individual contribution of the concentration of immigrants, firm size, economic activity, tenure, and level of qualification, that is, of the variables referring to employer and match characteristics. Therefore, Equation 8 summarizes the difference in the "Immigrant" coefficient between the two specifications considering Gelbach's decomposition method.

$$\hat{\delta} = \hat{\beta}_1^{specification\,2} - \hat{\beta}_1^{specification\,4} = \hat{\beta}_{14} \cdot \hat{\Gamma}_{concentration} + \hat{\beta}_{15} \cdot \hat{\Gamma}_{size} + \hat{\beta}_{16} \cdot \hat{\Gamma}_{industry} + \hat{\beta}_{17} \cdot \hat{\Gamma}_{tenure} + \hat{\beta}_{19} \cdot \hat{\Gamma}_{qualification}{}^{13} \qquad \text{(Eq. 8)}$$

It should be noted that Gelbach's method nests the traditional Oaxaca-Blinder decomposition technique, allowing for the use of Gelbach's auxiliary regression approach and variance formula to calculate Oaxaca-Blinder results. This is significant because many studies that utilize Oaxaca-Blinder do not mention sampling variances or standard errors. As a result, Gelbach's asymptotic variance formulas allow for the evaluation of statistical significance for the decomposition coefficients, addressing a constraint of the Oaxaca-Blinder technique (Gelbach, 2016).

The results of the method are presented in Annex F. Furthermore, similar to the findings of Carneiro *et al.* (2012), tenure and qualification levels account for 78.26% and 67.03% of the total variation in "Immigrant" for the female and male samples, respectively. Between the two variables, tenure contributes more to the modification, accounting for an increase of 0.0476 logarithmic points (39.34% of the difference) in the wages received by female immigrants and 0.0585 logarithmic points (40.34% of the difference) in the remuneration of male foreigners, representing approximately two and a half of the difference explained by the 27 covariates. On the other hand, the variation in qualification levels, measured by the four binary variables, explains 38.93% (women) and 26.69% (men) of the wage gap between immigrants and natives. Conversely, the firm size presents the smallest coefficients, explaining only 1.52% of the variation for women and 4.61% for men.

[13] "Industry" and "qualification" encompass all binary variables related, respectively, to economic activities (CAE) and qualification levels. Likewise, tenure includes the linear variable and its quadratic.

Subsequently, the decision was made to incorporate additional covariates in the decomposition equation to comprehend the impact of other factors, such as year, geographical location, and nationality, on the variable "Immigrant" and, consequently, on the wage disparity between natives and foreigners. Thus, by first adding dummy variables representing the years, it becomes possible to control for the effect of time and understand how each year of the sample influences the variation of the variable of interest.

It is evident in Table 23 of Annex F that tenure and qualification levels continue to exert the most substantial influence on the overall wage difference between the estimates of the "Immigrant" coefficient. They are responsible for 86.23% of the female total variation and 72.57% of the male. Conversely, the impact of the years on explaining the wage disparity is relatively low. For instance, in the female sample, 2018 accounts for only 0.29% of the wage difference compared to the base category, 2011.

Next, considering the addition of binary variables representing NUTS, Table 24 reveals that, concerning the base category (North), Lisbon accounts for the largest share of the coefficient variation among the regional dummies. For total wage divergence, women in Lisbon have a value of -9.12%, while men have a variation of -6.77%. This outcome suggests that, in the metropolitan area of the Portuguese capital, compared to other NUTS, especially Alentejo and Azores, there may be specific factors related to the region, such as labor supply and demand and cost of living, that play more relevant roles in determining the remuneration of foreign workers.

Finally, upon introducing dummy variables for nationality groups, the results presented in Table 25 reveal affiliation with the Brazil group, as compared to the reference category (EU14), exerts a considerable effect on the difference between the coefficient estimates of the variable "Immigrant" across specifications. Specifically, this effect amounts to -0.0591 logarithmic points for women (24.52% of the difference) and -0.0630 logarithmic points for men (18.92% of the difference). Similarly, men originating from the South Asian group also exert a noteworthy influence on the total salary difference concerning natives, accounting for approximately 13.96% of the wage divergence.

4.3 Accounting for unit fixed effects

As previously evidenced, through POLS, the sample observations were combined and estimated in "general" regressions, assuming, therefore, that the slope coefficients were constant between workers and firms. However, each employee, organization, and job may have specific characteristics, necessitating the accommodation of fixed effects (FE). Therefore, Specification 4 was re-estimated, but this time accounting for FE at the worker, firm, and qualification levels successively.

Before analyzing the results in Annex D, the Hausman test was conducted to determine whether the fixed effects or the random effects (RE) model was more appropriate. If the null hypothesis is rejected, it suggests that the FE model is preferred because the random effects may be correlated with one or more predictor variables (Gujarati & Porter, 2009). The outcomes showed a highly significant chi-squared p-value - below the 5% significance level - indicating that the null hypothesis was rejected and confirming the use of the FE model (see Annex C).

The results of applying worker-level fixed effects indicate similarity between the coefficients estimated under POLS with time-fixed effects quantitatively and qualitatively, not changing, therefore, the interpretation made for Specification 4. Consequently, it can be concluded that specific characteristics of workers do not significantly affect wages, and remuneration is mainly explained by individuals variations over time.

However, when considering the heterogeneity among companies, negative coefficients were obtained for the variable "Immigrant" (-0.0155 for women and -0.0139 for men), indicating that specific characteristics in each organization, such as management style, affect wages. Thus, under firm-level fixed effects, the remuneration of female immigrant workers upon arrival in Portugal is, on average, 1.54% lower than native workers with similar characteristics. For men, this value decreases to -1.38%. These results are consistent with those obtained by Carneiro *et al.* (2012), who found estimates of -2.55% and -4.38% for women and men, respectively, when applying establishment fixed-effects. Similarly, Damas de Matos (2017), concludes with a wage divergence of -8.60% between immigrants and Portuguese natives when accounting for firm heterogeneity.

Furthermore, there was a change in the signs of the YSM and its quadratic term coefficients. This inversion suggests that, initially, as immigrants assimilate into the country, their wages increase up to a turning point (19.72 years for women and 25.76 years for men), beyond which they start to experience less positive effects from the years of stay in Portugal.

The concentration variable also exhibited a reversal of signs compared to the fourth specification, indicating that an increase in the number of foreigners in the previous year positively impacts workers' salaries in the current year. More precisely, a 1% variation in the concentration of immigrants leads to a wage increase of 1.29% for women and 1.20% for men.

Lastly, incorporating fixed effects at the qualification level led to minimal coefficient adjustments, preserving their qualitative interpretation as observed when accounting for firm heterogeneity. Thus, the coefficients maintained the same signs, with only minor alterations in the values. “Immigrant”, for example, decreased to -1.51% for women and -1.33% for men.

4.4 Heterogeneity in nationality groups

Although the specifications conducted so far allow distinguishing between immigrant and native workers, they assume that the effects are homogeneous across different groups of foreign nationalities. However, as observed in the descriptive analysis, foreigners in the Portuguese private labor market exhibit important differences. For instance, EU14 immigrants have higher educational levels and receive above-average salaries. On the other hand, Chinese, African, and South Asian immigrants have a less significant proportion of workers with a bachelor's degree and occupy positions with lower remuneration. Therefore, it was decided to expand Specification 4 by adding binary variables representing nationalities, with EU14 as the reference category.

Moreover, the analysis of the fourth specification found that the concentration of immigrants in companies negatively affects remuneration. However, the regression does not allow for examining whether the relationship between salary and concentration changes depending on the workers' origin. To address this issue, multiplicative interactions between concentration of immigrants and the nationality binary variables were also included in the fourth specification.

$$\ln Wageh_{it} = \beta X_{it} + \delta_t + \gamma c_{it} + \theta m_{it} + \omega n_{it} + \varepsilon_{it}{}^{14} \qquad \text{(Spe. 5)}$$

Annex D presents the results for Specification 5, where it is possible to observe that female immigrants from Brazil receive approximately 15.21% less than workers from the European Union, all else being constant. Similarly, women from CEEC, China, PALOP, and SA also experience reductions of 15.04%, 14.01%, 14.70%, and 9.34% in remuneration compared to the base category. For men, the differences are -20.55% for Brazilians, -22.20% for Central and Eastern Europeans, -27.09% for Chinese, -23.51% for Africans, and -21.42% for South Asians. Therefore, it is evident that female Brazilian and male Chinese workers are the most negatively affected.

Furthermore, as the coefficient of "Immigrant" is positive for this specification, it is evident that the salary of EU14 immigrants with primary education, employed in manufacturing industries in the North, and assuming apprenticeship qualifications, was approximately 17.47% and 26.49% higher than that of native women and men, respectively, upon their arrival in Portugal in 2011.

Regarding the multiplicative terms, it is observed that China, for both genders, experiences the largest wage reduction among the groups when the concentration of foreigners in companies increases. This is interesting, given that this is the nationality group whose immigrants concentrate the most on companies with a foreign workforce. Cultural differences between Portugal and China and language divergences, may help explain why these workers prefer to work in places with other immigrants and are willing to accept less significant remuneration for it.

4.5 High-dimensionality regression

So far, new variables have been progressively added to the regression to explain workers' wages, reaching 63 independent terms by the fifth specification. However, Belloni *et al.* (2014) mention that applying standard estimation techniques, like OLS, to a large set of candidate predictors can lead the model to fit the noise in the data instead of the underlying pattern, resulting in poor performance on out-of-sample predictions. In other words, it may lead to a model with low bias and high variance, causing overfitting of the data and

[14] The vector n_{it} comprises the nationality dummies and their interaction with the concentration variable; while ω is a vector of coefficients that quantify the relationship between the variables in n_{it} and the wage.

inadequate generalization to new samples. As a consequence, overfitting can yield inflated regression coefficients, underestimated standard errors, small p-values, and large R^2 values compared to their population ones. Model interpretation relies on the accurate quantification of each of these properties, and it becomes increasingly misleading as the ratio of observations (n) to predictors (p) decreases (McNeish, 2015).

To avoid overfitting, it's necessary to find a balance between bias and variance. The trade-off between these two concepts refers to the dilemma in which unbiased estimates that minimize the total prediction error (as achieved by OLS) for a single seen dataset result in greater variation in the model's predictions across different samples. On the other hand, biased estimates are also not ideal, as even though they generalize better to new samples, they fail to capture the true relationship between predictor variables and the response term (McNeish, 2015).

In this sense, regularization is a technique employed to address the overfitting issue associated with OLS by introducing a penalty term into the process of minimizing the loss function (McNeish, 2015). Among the regularization methods, the Least Absolute Shrinkage and Selection Operator (LASSO) technique by Tibshirani (1996) sets a constraint on the absolute value of the sum of the regression coefficients, which will shrink coefficient estimates toward zero. The LASSO loss function thus can be written as an extension of OLS, incorporating an additional term to the sum of squared residuals. This term involves the sum of the absolute values of the coefficients betas multiplied by an adjustment parameter λ (see Equation 9).

The choice of λ is usually made through cross-validation, which randomly divides the dataset into k equal parts. While $k-1$ samples predict the model, the remaining sub-sample is used for validation. This process is performed k times, using each sub-sample once to validate the remaining ones. Thus, after completing the iterations and repeating this procedure for different values of λ, the parameter that minimizes the average prediction error is selected (Ranstam & Cook, 2018).

$$Minimize: \frac{1}{n}\sum_{i=1}^{n}(Y_i - X_i'\beta)^2 + \lambda\sum_{j=1}^{p}|\beta_j| \qquad \text{(Eq. 9)}$$

Effectively, by enabling coefficient estimates to be reduced to zero, LASSO, while obtaining coefficients with regularization, simultaneously performs variable selection, reducing the set of predictors to a more relevant group. Therefore, using Specification 5 as the base model, the method was applied to investigate the most relevant control variables that help to explain the wage disparity between immigrant and native workers within Portugal's private sector. This investigation was carried out by considering various subsets of the sample.

Nevertheless, the formulation of the fifth specification implicitly assumes that the coefficients of age, levels of education and qualification, year, NUTS, concentration, size, CAE, and tenure remain the same between foreigners and Portuguese. For instance, this assumption implies that for each additional year of age, immigrant and native workers experience the same incremental change in hourly wages. This theory can be tested including multiplicative variables resulting from the iteration between "Immigrant" and the other predictors of the model.

It can be observed in Table 26 of Annex G that for the complete sample (with 8,870,100 observations), out of the 104 variables, 103 were selected, with "Tenure^2 · Immi" being eliminated from the regression. For the female sample, on the other hand, the application of the technique indicated that 96 controls help to explain the wage disparity between foreign and native women, discarding, among other variables, the sectors of mining and electricity. Conversely, only three terms were removed from the male sample, resulting in 101 variables being retained.

In this regard, analyzing the coefficients obtained, it is evident that for all three samples (complete, female, and male), "Immigrant" remained positive and statistically significant, indicating that, for the omitted categories, foreigners have positive wage effects upon arrival in Portugal. More specifically, for the female sample, the coefficient increased to 0.172, and for the male, 0.565, meaning that European Union foreigners receive approximately 18.77% and 75.94% more than native women and men, respectively. For other immigrant nationalities, the wage penalty compared to EU14 remained.

Next, concerning age, for both genders, the signs of this variable and its squared term remained consistent with the previous results. For men, the statistically significant iteration between age and "Immigrant" suggests that an additional year of age increases the average wage of native workers by 2.49 log points, compared to a 0.82 log point increase for immigrants (2.49 – 1.67). Therefore, the age of male foreigners is worth less than one-third that of Portuguese workers, meaning that, for comparable employees within the same age group, an additional year of aging exacerbates the wage gap upon arrival. For women, on the other hand, an extra year of age generates an increase of 0.55 log points for natives and 0.64 log points (0.55 + 0.09) for immigrants. In this case, the aging of foreigners is more valued than that of Portuguese women.

Regarding the years since migration, the male sample showed again a negative sign for YSM and a positive sign for its squared term, indicating wage decline during assimilation. For women, on the other hand, YSM is positive, although very close to zero, suggesting a reward for female immigrants as their length of time in the country increases. Also, since YSM^2 was eliminated from the female regression, it can be concluded that there is a positive linear relationship between wages and years since migration for this gender.

Next, for educational qualifications, it is observed that all the interaction terms are negative, except for doctorate women. Therefore, the salary returns upon completing one more level of education (compared to having primary education) are lower for immigrants than for comparable nationals. Foreign women with doctorate stand out, with a wage increase 1.13 log points higher than that of Portuguese.

For the NUTS, the remuneration effects for female immigrants in Algarve and Centro, and male immigrants in Madeira, are more favorable (compared to North) than for national workers. However, in other regions, including Lisbon, foreign employees experience wage penalties compared to the Portuguese.

Concerning the years, immigration positively affects all coefficients except 2017 for the female sample. For instance, compared to 2011, in 2020, a male immigrant earned a salary of approximately 9.02 log points higher than that of a Portuguese worker with similar characteristics.

Regarding concentration, there is a persistence of negative coefficients for both genders, indicating detrimental effects on the pay of workers who are not nationals. Therefore, the negative impact on immigrant wages are 1.05 log points and 2.09 log points higher than that of Portuguese women and men, respectively.

On the other hand, in terms of firm size, foreign men and women receive positive rewards as the number of employees in the organization increases. Consequently, the impact on immigrant wages exceeds that of natives. This result may suggest that in larger companies, employees who speak multiple languages and have similarities with other cultures, as is often the case with many foreigners, represent advantages for the organization and are financially rewarded accordingly.

For economic activities, analyzing the results for the three sectors with the highest concentrations of immigrants in the sample, namely accommodation and food services, wholesale and retail trade, and administrative activities, it is observed that foreign workers are positively rewarded, and the wage penalty compared to manufacturing industries is smaller than that for nationals.

Regarding tenure, being an immigrant does not affect womens' wages, as the interactions between "Immigrant" and the variables "Tenure" and "$Tenure^2$" were considered irrelevant by LASSO. However, for male immigrants, the "Tenure · Immig" term is statistically significant, indicating a reduction in the positive impact that the number of years of tenure in a company has on these workers' wages.

Next, for qualification levels, the wage returns of moving up one more category (compared to being an apprentice) are higher for male immigrants than for comparable nationals. On the other hand, supervisors and highly skilled foreign women professionals are penalized compared to Portuguese natives.

Finally, continuing with the application of LASSO, aiming to understand how the formulated regression fits different geographical locations, the technique was applied to the seven NUTS within the sample. Naturally, the regional binary variables were eliminated from the specification.

Table 27 of Annex G shows that, as expected, Lisbon, being the largest sample, is the region where most controls are retained. Thus, out of 92 variables, 90 were considered valid for explaining the wage gap, with the iterations between "Immigrant" and "$Educ_5$" and "$Tenure^2$" being eliminated. On the other hand, for the Azores, the region with the smallest number of workers, 24 predictors were discarded, including the years 2012, 2017, and 2018. Finally, it is worth noting that, just like in the gender samples, the coefficient of the variable of interest ("Immigrant") remained positive for all NUTS, with Lisbon having the highest value (equal to 0.529).

4.6 Selection bias

As mentioned earlier, the QP does not provide information on the workers' country of birth, which can lead to a selectivity issue at the nationality level. This situation arises when the variable of interest affects the likelihood of individuals being included in the sample. It may introduce bias in the analysis because workers who were immigrants and later became naturalized citizens are now considered Portuguese and, therefore, are classified as nationals by the "Immigrant" variable. The outcome may result in distortion or underestimation of the actual effect of immigration on wages.

To verify if the sample is affected by this potential selection issue, Heckman's two-step procedure was applied. The method involves estimating two equations: a selection regression using a probit model with a binary variable as the dependent term; and an outcome equation associating the original dependent variable with the predictors and the inverse Mills ratio (IMR). The IMR is an additional term calculated using the estimated coefficients from the selection equation, and it allows for capturing the relationship between the data selection process and the endogenous variable (Verbeek, 2017).

More specifically, the selection equation estimates the probability of a worker being included in the sample (having "Immigrant" equal to 1) based on the predictor terms. Through this regression, the IMR is calculated, which corresponds to the ratio between the probability density function (PDF) and the cumulative distribution function (CDF) of the probit model. Then, the linear regression equation is estimated by OLS, using the dependent variable of interest (in this case, the logarithm of the hourly wage), the exogenous variables (including "Immigrant"), and the IMR correction term (Hoke & Bendig, 2022).

The significance level and sign of the IMR's coefficient indicate the correlation between the error terms of the selection equation and the outcome equation. Therefore, a significantly positive beta coefficient suggests that unobserved factors positively affect the estimated relationship (Hoke & Bendig, 2022).

Annex H reveals that the IMR is statistically significant at the 1% level, signaling potential selectivity issues: individuals categorized as immigrants tend to exhibit a propensity for receiving lower wages compared to those in the base category. In essence, this suggests the presence of a wage gap between natives and immigrants.

4.7 Conclusion

Initially, three regressions were estimated using Pooled OLS with time fixed effects for the logarithm of hourly wages, each including the variable of interest of the study ("Immigrant"). With Specification 2, when exclusively controlling workers' characteristics, consistent with many regressions formulated in the literature, immigrants arriving in

Portugal received lower wages than natives. In Specification 3, adding variables related to firm characteristics, the wage penalty for immigrants persisted in the female sample. However, in Specification 4, after accounting for match characteristics between employers and workers, such as tenure and qualification levels, a notable sign reversal was observed in the coefficient for "Immigrant" for both genders. Therefore, including match variables indicates that, initially, such factors play a relevant role in generating higher wages for foreign workers upon their arrival in the country. Confirmation of this theory was achieved through the application of Gelbach's decomposition, which identified match characteristics as the primary drivers of the observed variation in "Immigrant" across specifications.

Adding other binary variables to Gelbach's decomposition equation allowed understand that the nationality group to which the immigrant belongs affects the wage difference between foreigners and natives more significantly than factors related to companies (concentration, size, and industry), years, and NUTS.

Next, the application of fixed effects revealed that worker-specific characteristics do not significantly affect wages. On the other hand, the addition of FE at the firm and qualification levels indicated a wage penalty for foreign workers. Thus, there is systemic heterogeneity among firms and among job positions that help explain the lower wages of immigrants compared to nationals.

Subsequently, binary terms representing nationality groups and interaction variables between these dummies and “concentration” were added to Specification 4. This addition confirmed the results of the descriptive analysis, indicating higher wages for workers from the European Union. Thus, all other foreign nationalities, especially female workers from Brazil and male immigrants from China, are penalized compared to the base category (EU14). Additionally, it is among the Chinese sample where the largest wage reduction is observed for increases in foreign worker concentration in firms.

Upon applying LASSO, it was observed that there was no significant reduction in dimensionality for the samples, indicating that the number of regressors in the specifications remained largely unchanged. Several reasons could lead to this outcome: low multicollinearity, strong relationships with the dependent variable, small effect size for dropped variable, moderate regularization strength, complexity of the true model, sample size, and low noise in the data. Notably, in datasets with lower noise levels, LASSO can effectively estimate coefficients and identify relevant variables. The lower noise contributes to a reduced likelihood of discarding variables that carry some true signal, even if their impact may be less pronounced compared to others.

Also, the specification used in applying the regularization method considered the interaction between "Immigrant" and other predictor variables of the model, providing some relevant insights. Among these, the estimated returns for education (compared to having only primary qualification) for immigrants are lower than for natives across all categories (except women with doctorate). Consequently, on average, the salaries of foreigners with higher formal education are more penalized in the Portuguese private sector, suggesting imperfect human capital portability between countries (Cabral & Duarte, 2012).

Finally, applying Heckman's two-step procedure indicated selectivity problems within the sample, suggesting wage disparities between immigrants and Portuguese.

5 Conclusion

Over the past few decades, Portugal has undergone significant transformations concerning its role in international migrations. The country, which until the 1980s experienced moderate immigration flows, witnessed, for the first time by the end of the 20th century, a positive migratory balance due to economic expansion driven by its entry into the European Community. However, even with the acceleration in the influx of new immigrants, the proportion of foreigners in the country remains relatively low compared to other European Union member states. Nonetheless, it is relevant to highlight the contribution of this group in bolstering the active workforce and mitigating the effects of the demographic aging of the Portuguese population. Moreover, initiatives in entrepreneurship, capital investments, and payments to the Social Security system reinforce the role of immigrants in Portugal.

Despite the mentioned contributions, the Portuguese labor market still displays segmentation based on nationality, with foreign workers often concentrated in less skilled and lower-paid activities and economic sectors. However, diversity is observed among the resident foreign population, as the demographic structure of immigrants is not homogeneous. As a result, immigrant groups with recognized qualifications and wages above the national average can be identified.

Therefore, given the increase in immigration flows and the diversity of immigrants within Portuguese territory, it becomes pertinent to analyze the relative evolution of wages among workers. Understanding the factors influencing remuneration is essential for formulating policies and programs aimed at the successful integration of immigrants into the country while also fostering a more equitable labor market.

To achieve this, the Quadros de Pessoal database was employed, a longitudinal dataset containing worker-firm pairs' records from the period 2011 to 2020. This enabled the analysis of workers' wages and the evaluation of whether the remuneration difference between natives and immigrants is due to firm characteristics or the combinations they form with individuals. This analysis was conducted by progressively including new variables in the regression, followed by applying the decomposition method proposed by Gelbach (2016), and finally employing LASSO, a regularization technique aimed at selecting relevant variables for the regression.

Through these approaches, the persistence of a wage differential between native and immigrant workers was observed, with this discrepancy being sensitive to the set of applied regressors. Including match characteristics, such as tenure and qualification levels, and nationality dummies was the primary factor contributing to the increase in divergence.

Furthermore, it is emphasized that among the analyzed nationality groups (Brazil, CEEC, China, EU14, PALOP, and SA), immigrants from the European Union are primarily responsible for the assessment of higher wages, being the group with the highest educational qualifications. On the other hand, China, besides being one of the most negatively penalized nationalities, experiences the largest wage reduction in the sample as the concentration of foreigners in firms increases.

Additionally, despite not resulting in a significant decrease in dimensionality, the LASSO estimation approach is a systematic and helpful tool for variable selection that is not yet widely used in the literature on migration economics, especially in Portugal.

In summary, this study highlighted that the endurance of the wage gap among workers in Portugal is a multifaceted issue influenced by a range of economic factors.

One significant aspect is skill mismatches. Immigrants often possess skill levels and qualifications that differ from those of native workers. Unfortunately, these skills might not be fully recognized or valued within the Portuguese labor market, resulting in lower wages for foreign workers. On the other hand, cases such as those of immigrants from the European Union also show that nationality has an impact on the recognition of qualifications, which, in this case, are more valued than those of native workers.

Occupational segregation is another contributor to the wage gap. Immigrants tend to find themselves clustered in specific industries or sectors (like accommodation and food service, and wholesale and retail trade) due to factors such as social networks, discrimination, or difficulties in having foreign qualifications recognized. Unfortunately, if these sectors generally offer lower wages on average, it perpetuates the wage disparity.

Discrimination and bias also play a significant role. Although the focus of this work was not on measuring wage discrimination against migrants, it is relevant to mention that prejudice against the group can hinder their access to better job opportunities and promotions, ultimately resulting in lower wages compared to equally qualified native workers. This bias can manifest in various ways, such as biased recruitment practices or disparities in promotions.

Educational attainment is another contributing factor. Disparities in education between native and immigrant populations can lead to wage differences. Immigrants with lower levels of educational attainment might face limited opportunities for higher-paying jobs, thus impacting their earning potential.

Cyclical labor market dynamics is also relevant. Economic downturns or fluctuations in labor demand can disproportionately affect immigrants, leading to job losses or reduced wages. These factors can hinder the economic progression of immigrant workers and contribute to wage disparities.

Furthermore, there are additional factors not covered in this dissertation that can influence the wage gap and be analyzed in future research. Institutional barriers, such as immigration policies and labor market regulations, can either support or hinder immigrant integration into the workforce. Remittances and motivations related to job acceptance might also play a role. Language barriers can impact job prospects, as fluency in the local language is essential for successful labor market integration. Additionally, the lack of access to professional networks can limit immigrants' opportunities and affect their wage outcomes.

Finally, further research into the wage gap among immigrants in Portugal could encompass a broader range of methodologies, including other regularization approaches, to assess the robustness of these findings. Additionally, there is potential to explore specific factors such as nationalities, gender, and regional and economic activity variations. For instance, the study of the immigrant wage gap between genders remains limited within the country. Also, one of the outcomes of this study was the observed penalized salary in the Northern region. Consequently, research focused on this economically significant Portuguese region would yield valuable insights.

References

ACM. (2022). Lei de Estrangeiros: O que mudou?. Alto Comissariado para as migrações. Retrieved June 20, 2023, from https://www.acm.gov.pt/pt/acm.

Aldashev, A.; Gernandt, J.; & Thomsen, S. L. (2012). The immigrant-native wage gap in Germany. *Jahrbücher für Nationalökonomie und Statistik*, *232* (5), 490-517.

Andersson, F.; Garcia-Perez, M.; Haltiwanger, J.; McCue, K.; & Sanders, S. (2014). Workplace concentration of immigrants. *Demography*, *51* (6), 2281-2306.

AT. (2016). Regime Fiscal Para o Residente Não Habitual. Autoridade Tributária e Aduaneira. Retrieved June 20, 2023, from https://portaldascomunidades.mne.gov.pt/pt/.

Baganha, M. I.; Marques, J. C.; & Góis, P. (2004). Novas migrações, novos desafios: A imigração do Leste Europeu. *Revista Crítica de Ciências Sociais* (69), 95-115. https://doi.org/10.4000/rccs.1340

Basso, G.; & Peri, G. (2020). Internal Mobility: The Greater Responsiveness of Foreign-Born to Economic Conditions. *Journal of Economic Perspectives*, *34* (3), 77-98.

Beach, C. M.; & Worswick, C. (1993). Is There a Double-Negative Effect on the Earnings of Immigrant Women? *Canadian Public Policy/Analyse de Politiques*, *19*, 36-53. https://doi.org/10.2307/3551789

Becker, G. S. (1957). The economics of discrimination. University of Chicago Press. https://archive.org/.

Belloni, A.; Chernozhukov, V.; & Hansen, C. (2014). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, *28* (2), 29-50.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, *8*, 436-455. https://doi.org/10.2307/144855

Borjas, G. J. (1982). The earnings of male Hispanic immigrants in the United States. *ILR Review*, *35* (3), 343-353. https://doi.org/10.2307/2522814

Borjas, G. J. (1985). Assimilation, changes in cohort quality, and the earnings of immigrants. *Journal of Labor Economics*, *3* (4), 463-489. https://doi.org/10.1086/298065

Borjas, G. J. (1994). The Economics of Immigration. *Journal of Economic Literature*, *32*, 1667–1717. http://www.jstor.org/stable/2728791

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly journal of economics*, *118* (4), 1335-1374.

Boyd, M. (1984). At a Disadvantage: The Occupational Attainments of Foreign Born Women in Canada. *International migration review*, *18* (4), 1091-1119. https://doi.org/10.2307/2546074

Boyd, M.; & Grieco, E. (2003). Women and migration: Incorporating gender into international migration theory.

Cabral, S.; & Duarte, C. (2012). O Diferencial de Salários dos Imigrantes no Mercado de Trabalho Português. *Boletim Económico do Banco de Portugal*, 85-103.

Cabral, S.; & Duarte, C. (2013). Mind the gap! The relative wages of immigrants in the Portuguese labour market. In *Working Paper 05-2013, Banco de Portugal, Lisboa, Portugal.*

Cabral, S.; & Duarte, C. (2016). Lost in Translation? The Relative Wages of Immigrants in the Portuguese Labour Market. *International Review of Applied Economics*, *30* (1), 27-47.

Card, D. (2005). Is The New Immigration Really So Bad? *The economic journal*, *115* (507), F300-F323.

Carliner, G. (1980). Wages, earnings and hours of first, second, and third generation American males. *Economic Inquiry*, *18* (1), 87-102. https://doi.org/10.1111/j.1465-7295.1980.tb00561.x

Carneiro, A.; Fortuna, N.; & Varejão, J. (2012). Immigrants at new destinations: how they fare and why. *Journal of Population Economics*, *25*, 1165-1185.

Carneiro, A. (2021). Decomposition Methods in Economics to Assess which Covariates Matter. Retrieved June 20, 2023, from https://sigarra.up.pt/fep/en/web_page.Inicial.

Chiswick, B. R. (1978). The Effect of Americanization on the Earnings of Foreign-Born Men. In *Interdisciplinary Perspectives on the New Immigration* (pp. 111-136). Routledge.

Csonto, B.; Liu, L.; Suphaphiphat, N.; & Westelius, N. (2015). International Migration: Recent Trends, Economic Impacts, And Policy Implications.

Damas de Matos, A. (2017). Firm heterogeneity and immigrant wage assimilation. *Applied Economics Letters*, *24* (9), 653-657.

De Brauw, A. (2017). Does Immigration Reduce Wages. *Cato Journal*, *37*, 473-480.

Dhrymes, P. J. (1978). Introductory econometrics. Springer-Verlag. https://archive.org/.

Duleep, H. O.; & Regets, M. C. (1999). Immigrants and human-capital investment. *American Economic Review*, *89*, 186-191.

Eurostat. (2023a). Monthly minimum wages - bi-annual data (2018 - 2023). European Statistical Office. Retrieved April 10, 2023, from https://ec.europa.eu/eurostat/en/.

Eurostat. (2023b). Immigration by age and sex. European Statistical Office. Retrieved May 16, 2023, from https://ec.europa.eu/eurostat/en/.

Eurostat. (2023c). Emigration by age and sex. European Statistical Office. Retrieved May 16, 2023, from https://ec.europa.eu/eurostat/en/.

Eurostat. (2023d). Population structure indicators at national level (1960 - 2022). European

Statistical Office. Retrieved April 10, 2023, from https://ec.europa.eu/eurostat/en/.

FFMS. (2023). População estrangeira com estatuto legal de residente: total e por algumas nacionalidades. PORDATA – Estatísticas, gráficos e indicadores de Municípios, Portugal e Europa. Retrieved June 14, 2023, from https://www.pordata.pt.

Freijeiro-González, L.; Febrero-Bande, M.; & González-Manteiga, W. (2022). A critical review of LASSO and its derivatives for variable selection under dependence among covariates. *International Statistical Review*, *90* (1), 118-145.

Friedberg, R. M. (2000). You can't take it with you? Immigrant assimilation and the portability of human capital. *Journal of Labor Economics*, *18*, 221-251. https://doi.org/10.1086/209957

Gaspar, S. (2018). Percursos migratórios e trajetórias de vida de descendentes de imigrantes chineses. *Sociologia, Problemas e Práticas* (87), 109-127.

Gelbach, J. B. (2016). When do covariates matter? And which ones, and how much? *Journal of Labor Economics*, *34* (2), 509-543.

Góis, P.; Abecasis, F.; Alfaiate, J.; Hancock, M.; & Lozano, L. (2019). Migrações e Desenvolvimento em Portugal - Avançar nas Práticas: Rumo à Inclusão e Coesão Social.

Góis, P.; & Marques, J. C. (2007). Estudo Prospectivo Sobre Imigrantes Qualificados em Portugal. Observatório da Imigração, ACIDI, IP.

Góis, P.; Marques, J. C.; Valadas, C.; Leite, A.; & Nolasco, C. (2018). Discriminação no Recrutamento e Acesso ao Mercado de Trabalho de Imigrantes e Portugueses de Origem Estrangeira (Vol. 63). Observatório das Migrações, ACM, IP.

Goldin, I.; Pitt, A.; Nabarro, B.; & Boyle, K. (2018). Migration and the Economy: Economic Realities, Social Impacts & Political Choices.

Gomes, N. (2009). Os portugueses na Venezuela. *Relações Internacionais*, 83-92.

Gordon, M. M. (1964). Assimilation in American life: The role of race, religion, and national origins. Oxford University Press, USA. https://archive.org/.

Gujarati, D. (2014). Econometrics by example. Bloomsbury Publishing.

Gujarati, D. N.; & Porter, D. C. (2009). Basic econometrics (5th ed.). Prentice Hall.

Hoke, J.; & Bendig, D. (2022). Selection bias and Heckman two-stage estimation. Retrieved April 10, 2023, from https://www.statisticslab.org/.

Hunt, J.; & Gauthier-Loiselle, M. (2010). How Much Does Immigration Boost Innovation? *American Economic Journal: Macroeconomics*, *2* (2), 31-56.

IAPMEI. (2022). StartUP Visa. Agência para a Competitividade e Inovação. Retrieved June 20, 2023, from https://www.iapmei.pt/.

ILO. (2020). The migrant pay gap: Understanding wage differences between migrants and nationals.

ILO. (2021). ILO Global Estimates on International Migrant Workers: Results and Methodology. Retrieved June 22, 2023, from https://www.ilo.org/.

INE. (2023). Índice de preços no consumidor (IPC, Base - 2012) por Localização geográfica e Consumo individual por objectivo. Portal do INE. Retrieved April 10, 2023, from www.ine.pt.

IOM. (2021). World Migration Report 2022. Geneva, Switzerland: International Organization for Migration.

Iorio, J. C.; & Ferreira, S. d. A. (2013). Fluxos migratórios de brasileiros em portugal: o retorno e a "nova vaga dos em vias de qualificação". *Leopoldianum*, *39* (107-9), 31-48.

Lemieux, T. (2006). The "Mincer equation" thirty years after schooling, experience, and earnings. In *Jacob Mincer A Pioneer of Modern Labor Economics*. Springer, Boston, MA. https://doi.org/10.1007/0-387-29175-X_11

Levy, N.; Pisarevskaya, A.; & Scholten, P. (2020). Between fragmentation and institutionalisation: the rise of migration studies as a research field. *Comparative Migration Studies*, *8* (1), 1-24.

Long, J. E. (1980). The Effect of Americanization on Earnings: Some Evidence for Women. *Journal of Political Economy*, *88* (3), 620-629. https://doi.org/10.1086/260892

Longva, P.; & Raaum, O. (2003). Earnings assimilation of immigrants in Norway–A reappraisal. *Journal of Population Economics*, *16* (1), 177-193.

Lutz, H. (2010). Gender in the migratory process. *Journal of ethnic and migration studies*, *36* (10), 1647-1663. https://doi.org/10.1080/1369183X.2010.489373

Malheiros, J.; & Esteves, A. (2013). Diagnóstico da população imigrante em Portugal: desafios e potencialidades (989685050X).

McNeish, D. M. (2015). Using lasso for predictor selection and to assuage overfitting: A method long overlooked in behavioral sciences. *Multivariate behavioral research*, *50* (5), 471-484.

Mincer, J. (1974). Schooling, Experience, and Earnings. Human Behavior & Social Institutions No. 2. National Bureau of Economic Research.

Mota, P. R. (2017). Austeridade Expansionista-Como Matar uma Ideia Zombie. Leya.

Nicodemo, C.; & Ramos, R. (2012). Wage Differentials Between Native and Immigrant Women in Spain: Accounting for Differences in Support. *International Journal of Manpower*.

Nielsen, H. S.; Rosholm, M.; Smith, N.; & Husted, L. (2004). Qualifications, discrimination, or assimilation? An extended framework for analysing immigrant wage gaps. *Empirical Economics*, *29* (4), 855-883. https://doi.org/10.1007/s00181-004-0221-9

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International economic review*, *14*, 693-709. https://doi.org/10.2307/2525981

OECD. (2023a). Native-born unemployment (indicator). . Organisation for Economic Co-operation and Development. Retrieved July 04, 2023, from https://doi.org/10.1787/0f9d8842-en.

OECD. (2023b). Foreign-born unemployment (indicator). . Organisation for Economic Co-operation and Development Retrieved July 04, 2023, from https://doi.org/10.1787/ba5d2ce0-en.

Oliveira, C. R.; & Gomes, N. (2014). Monitorizar a integração de imigrantes em Portugal: relatório estatístico decenal. Lisbon.

Oliveira, C. R. (2021). Indicadores de Integração de Imigrantes: Relatório Estatístico Anual 2021. Observatório das Migrações.

Pinho, A. (2013). A evolução das políticas de imigração e asilo em Portugal no contexto de uma Europa Comunitária. *População e Sociedade.*, 123.

Portugal. (2009). Código do Trabalho [Lei nº 7/2009]. Diário da República n.º 30/2009, Série I de 2009-02-12.

Portugal. (2018). Portaria 328/2018, de 19 de Dezembro. Diário da República n.º 244/2018, Série I de 2018-12-19.

Portugal. (2020). Resolução do Conselho de Ministros n.º 16/2020. Diário da República n.º 62/2020, Série I de 2020-03-27.

Portugal. (2021). Diário da República - Decreto n.º 27/2021.

Portugal. (2023). Estatutos das ordens profissionais – saiba o que muda. Retrieved August 02, 2023, from https://www.portugal.gov.pt/pt/gc23.

Ranstam, J.; & Cook, J. (2018). LASSO regression. *Journal of British Surgery*, *105* (10), 1348-1348.

Sampaio, V. R. B. d. (2017). Os processos de regularização de imigrantes realizados em Portugal: que avaliação?

Sanromá, E.; Ramos, R.; & Simón, H. (2015). Portability of Human Capital and Immigrant Overeducation in Spain. *Population Research and Policy Review*, *34* (2), 223-241.

Schoeni, R. F. (1997). New evidence on the economic progress of foreign-born men in the 1970s and 1980s. *Journal of Human resources*, 683-740. https://doi.org/10.2307/146426

Scholten, P.; Pisarevskaya, A.; & Levy, N. (2022). An introduction to migration studies: the rise and coming of age of a research field. In *Introduction to Migration Studies: An Interactive Guide to the Literatures on Migration and Diversity* (pp. 3-24). Springer.

SEF. (2012). ARI - Residence permit for investment activity. Serviço de Estrangeiros e

Fronteiras. Retrieved June 22, 2023, from https://www.sef.pt/.

Shamsuddin, A. F. (1998). The double-negative effect on the earnings of foreign-born females in Canada. *Applied Economics*, *30*, 1187-1201. https://doi.org/10.1080/000368498325084

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58* (1), 267-288.

Tilcsik, A. (2021). Statistical discrimination and the rationalization of stereotypes. *American sociological review*, *86* (1), 93-122.

UN/DESA. (2020a). International Migration 2020 Highlights. United Nations Department of Economic and Social Affairs. Retrieved April 10, 2023, from https://www.un.org/.

UN/DESA. (2020b). World Population Policies 2019. United Nations Department of Economic and Social Affairs. Retrieved April 10, 2023, from https://www.un.org/.

UNHCR. (2023a). Refugee Data Finder. The United Nations High Commissioner for Refugees. Retrieved April 10, 2023, from https://www.unhcr.org/.

UNHCR. (2023b). Ukraine Refugee Situation. The United Nations High Commissioner for Refugees. Retrieved April 12, 2023, from https://data2.unhcr.org/en/situations.

United Nations Educational, S. a. C. O. U. (2018). Global Education Monitoring Report 2019: Migration, Displacement and Education – Building Bridges, not Walls. Paris.

UNODC. (2022). Global Report on Trafficking in Persons 2022. United Nations Office on Drugs and Crime. Retrieved June 14, 2023, from https://www.un-ilibrary.org/.

Verbeek, M. (2017). A guide to modern econometrics. John Wiley & Sons.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the econometric society*, *48*, 817-838. https://doi.org/10.2307/1912934

Yavuz, C. I.; Güler, Ç.; Eryurt, M.; & Vaizoğlu, S. A. (2020). Changing of risk factors related to diarrhoea among children aged under 5 within ten years in Turkey. *Central European journal of public health*, *28*.

Annexes

Annex A - Variables definition

Table 12 - Description of the variables of interest

Variables	Description
ln Wageh	Natural logarithm of the ratio between the real monthly base wage and the total normal hours of work
Immigrant	Dummy variable for worker's country of origin:
	= 1 if immigrant
	= 0 if national worker
YSM	Proxy for number of years an immigrant worker has resided in Portugal (in years). For native workers, this variable assumes value of zero.
YSM^2	Years since migration squared (in years)
Age	Worker's age (in years)
Age^2	Age squared (in years)
Tenure	Number of years with the current employer (in years)
$Tenure^2$	Tenure squared (in years)
Concentration	Percentage of immigrant workers in the firm (lagged by one year, in logs)
Size	Number of employees in the firm (in logs)
Education levels	Dummy variables for worker's education levels:
$Educ_1$	= 1, if primary is the highest level the worker completed (omitted category)
$Educ_2$	= 1, if secondary is the highest level the worker completed
$Educ_3$	= 1, if college is the highest level the worker completed
$Educ_4$	= 1, if master is the highest level the worker completed
$Educ_5$	= 1, if doctorate is the highest level the worker completed
Qualification levels	Dummy variables for worker's qualification levels:
$Qual_1$	= 1, if Top or intermediary executives
$Qual_2$	= 1, if Supervisors
$Qual_3$	= 1, if Highly skilled or skilled
$Qual_4$	= 1, if Semi-skilled or unskilled
$Qual_5$	= 1, if Apprentices (omitted category)
NUTS II	Dummy variables for the firm's location:
North	= 1, if NUTS II North (omitted category)
Algarve	= 1, if NUTS II Agarve
Centre	= 1, if NUTS II Centre
Lisbon	= 1, if NUTS II Lisbon
Alentejo	= 1, if NUTS II Alentejo
Azores	= 1, if NUTS II Azores
Madeira	= 1, if NUTS II Madeira

CAE		Dummy variables for the firm's economic activity:
	CAE_1	= 1, if Agriculture, farming of animals, hunting and forestry
	CAE_2	= 1, if Mining and quarrying
	CAE_3	= 1, if Manufacturing (omitted category)
	CAE_4	= 1, if Electricity, gas, steam, cold and hot water and cold air
	CAE_5	= 1, if Water collection, treatment and distribution; sewerage, waste management and remediation activities
	CAE_6	= 1, if Construction
	CAE_7	= 1, if Wholesale and retail trade; repair of motor vehicles and motorcycles
	CAE_8	= 1, if Transportation and storage
	CAE_9	= 1, if Accommodation and food service activities
	CAE_{10}	= 1, if Information and communication activities
	CAE_{11}	= 1, if Financial and insurance activities
	CAE_{12}	= 1, if Real estate activities
	CAE_{13}	= 1, if Consultancy, scientific and technical activities
	CAE_{14}	= 1, if Administrative and support service activities
	CAE_{15}	= 1, if Public administration and defence; compulsory social security
	CAE_{16}	= 1, if Education
	CAE_{17}	= 1, if Human health and social work activities
	CAE_{18}	= 1, if Arts, entertainment, sports and recreation activities
	CAE_{19}	= 1, if Other service activities
	CAE_{20}	= 1, if Activities of households as employers; undifferentiated goods and services producing activities of households for own use
	CAE_{21}	= 1, if Activities of extraterritorial organizations and bodies
Year		Dummy variables for year:
	$Year_{11}$	= 1, if year is 2011 (omitted category)
	$Year_{12}$	= 1, if year is 2012
	$Year_{13}$	= 1, if year is 2013
	$Year_{14}$	= 1, if year is 2014
	$Year_{15}$	= 1, if year is 2015
	$Year_{16}$	= 1, if year is 2016
	$Year_{17}$	= 1, if year is 2017
	$Year_{18}$	= 1, if year is 2018
	$Year_{19}$	= 1, if year is 2019
	$Year_{20}$	= 1, if year is 2020
Nationality		Dummy variables for worker's group of nationality:
	Brazil	= 1, if Brazil
	CEEC	= 1, if Czech Republic, Estonia, Hungary, Latvia, Lithuania, Moldova, Poland, Romania, Russia, Serbia, Slovakia, Slovenia or Ukraine
	China	= 1, if China
	EU14	= 1, if Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Spain, Sweden or United Kingdom (omitted category)
	PALOP	= 1, if Angola, Cape Verde, Guinea-Bissau, Mozambique or São Tomé and Príncipe
	Portugal	= 1, if Portugal
	SA	= 1, if Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan or Sri Lanka
	Others	= 1, if Other nationalities

Source: Own elaboration.

Annex B - Descriptive statistics

Table 13 - Evolution of the real minimum wage (in €) and CPI (base 2012) in Portugal

Year	Minimum Wage	CPI
2011	485	97.302
2012	485	100.000
2013	485	100.274
2014	485	99.996
2015	505	100.483
2016	530	101.094
2017	557	102.477
2018	580	103.496
2019	600	103.846
2020	635	103.833

Source: FFMS & INE. Own elaboration.

Table 14 - Mean of the variables for the female sample (2011 - 2020)

	Portugal	Brazil	CEEC	China	EU14	PALOP	SA	Others
Years since migration (in years)	-	5.12	7.10	6.03	6.75	8.90	2.07	6.11
Age (in years)	39.71	36.69	39.05	36.43	36.87	37.74	32.38	38.09
Tenure (in years)	9.30	2.78	4.33	2.76	4.41	5.16	1.22	3.64
Concentration of immigrants	3.80%	24.65%	24.04%	62.84%	23.14%	15.70%	41.61%	23.82%
EDUCATION LEVELS								
Primary	39.14%	49.37%	53.06%	82.94%	14.82%	65.03%	63.48%	40.56%
Secondary	29.85%	38.74%	31.95%	11.37%	39.30%	26.70%	30.42%	32.08%
College	27.88%	10.81%	12.84%	4.29%	38.70%	7.57%	4.85%	22.97%
Mater	2.78%	0.90%	1.87%	1.17%	5.64%	0.63%	0.91%	3.56%
Doctorate	0.36%	0.19%	0.28%	0.23%	1.53%	0.07%	0.35%	0.82%
QUALIFICATION LEVELS								
Top or intermediary executives	18.02%	4.75%	4.86%	6.47%	23.79%	4.65%	2.48%	12.18%
Supervisors	16.22%	5.86%	5.45%	6.55%	18.21%	6.06%	3.56%	9.02%
Highly skilled or skilled	55.35%	65.16%	65.74%	68.85%	52.69%	60.22%	64.63%	60.87%
Semi-skilled or unskilled	8.31%	17.97%	19.74%	4.67%	3.32%	24.33%	21.16%	14.46%
Apprentices	2.10%	6.26%	4.20%	13.47%	1.99%	4.75%	8.17%	3.47%
NUTS II								
North	24.52%	12.24%	10.33%	23.04%	17.34%	8.01%	3.06%	14.38%
Algarve	3.73%	9.39%	21.85%	5.69%	11.92%	7.01%	11.37%	12.11%
Centre	13.98%	12.02%	19.02%	12.23%	7.05%	6.55%	12.72%	14.79%
Lisbon	51.46%	61.65%	39.09%	49.83%	59.30%	75.39%	60.71%	45.48%
Alentejo	3.71%	3.86%	8.44%	7.31%	2.79%	2.32%	11.76%	8.92%
Azores	1.03%	0.30%	0.16%	0.94%	0.39%	0.34%	0.13%	0.93%
Madeira	1.57%	0.54%	1.11%	0.97%	1.20%	0.37%	0.25%	3.39%
CAE								
CAE_1	0.88%	1.48%	10.10%	0.02%	0.93%	0.76%	13.87%	11.09%
CAE_2	0.08%	0.02%	0.08%	0.00%	0.05%	0.04%	0.00%	0.08%
CAE_3	19.72%	9.99%	17.29%	0.58%	11.12%	6.46%	5.87%	11.34%
CAE_4	0.20%	0.01%	0.01%	0.00%	0.10%	0.07%	0.00%	0.03%
CAE_5	0.55%	0.25%	0.85%	0.00%	0.14%	0.52%	0.31%	0.52%
CAE_6	1.18%	1.33%	1.58%	0.15%	1.04%	1.04%	0.45%	1.07%
CAE_7	17.68%	16.00%	12.73%	68.91%	8.92%	13.31%	13.49%	12.11%
CAE_8	3.29%	1.25%	1.61%	0.23%	2.05%	1.04%	0.11%	0.99%
CAE_9	7.94%	28.31%	25.08%	23.55%	7.85%	31.42%	50.20%	17.62%
CAE_{10}	3.54%	2.02%	1.87%	0.50%	6.18%	1.02%	0.91%	3.62%
CAE_{11}	6.71%	1.07%	0.88%	0.97%	3.41%	1.54%	0.31%	2.00%
CAE_{12}	0.38%	1.20%	1.88%	2.15%	1.90%	0.82%	1.19%	1.34%
CAE_{13}	3.10%	2.58%	2.38%	1.21%	7.15%	2.03%	0.78%	4.19%
CAE_{14}	7.95%	11.80%	9.55%	0.99%	29.94%	13.92%	7.56%	15.50%
CAE_{15}	0.23%	0.08%	0.08%	0.00%	0.13%	0.12%	0.00%	0.15%

CAE_{16}	2.25%	1.09%	0.88%	0.35%	5.35%	2.08%	0.88%	3.23%
CAE_{17}	21.15%	15.46%	8.69%	0.16%	9.81%	19.56%	1.84%	11.07%
CAE_{18}	0.73%	0.96%	1.17%	0.19%	1.16%	0.80%	0.25%	1.27%
CAE_{19}	2.42%	5.12%	3.26%	0.05%	2.78%	3.43%	1.84%	2.76%
CAE_{20}	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
CAE_{21}	0.00%	0.00%	0.01%	0.00%	0.01%	0.03%	0.14%	0.01%
FIRM SIZE								
Micro	2.62%	24.37%	24.30%	72.49%	12.23%	14.01%	23.50%	17.00%
Small	12.45%	28.78%	30.82%	24.58%	14.18%	24.78%	36.20%	23.04%
Medium	27.10%	20.91%	24.28%	1.46%	20.75%	23.38%	19.73%	25.98%
Large	57.83%	25.94%	20.60%	1.47%	52.83%	37.83%	20.56%	33.98%

Notes: Micro (less than 10 employees), small (10 to 49 employees), medium (50 to 249 employees) and large (more than 250 employees).

Source: Quadros de Pessoal. Own elaboration.

Table 15 - Mean of the variables for the male sample (2011 - 2020)

	Portugal	Brazil	CEEC	China	EU14	PALOP	SA	Others
Years since migration (in years)	-	5.30	8.69	6.20	6.37	9.13	2.30	6.14
Age (in years)	40.24	35.60	41.15	36.69	38.35	38.09	32.50	38.12
Tenure (in years)	9.44	2.70	4.70	2.41	4.27	4.70	1.02	3.53
Concentration of immigrants	3.98%	26.36%	26.17%	64.88%	22.92%	17.88%	46.27%	25.26%
EDUCATION LEVELS								
Primary	48.57%	55.81%	64.63%	85.33%	18.48%	67.19%	71.02%	53.81%
Secondary	30.26%	33.57%	29.48%	11.31%	38.96%	25.90%	24.83%	29.47%
College	18.38%	9.48%	5.25%	2.64%	35.06%	6.41%	3.53%	14.01%
Mater	2.52%	0.93%	0.43%	0.44%	5.19%	0.46%	0.48%	1.97%
Doctorate	0.27%	0.21%	0.21%	0.27%	2.32%	0.03%	0.14%	0.74%
QUALIFICATION LEVELS								
Top or intermediary executives	17.94%	7.69%	2.51%	6.80%	31.87%	5.53%	1.65%	10.33%
Supervisors	17.47%	7.67%	4.28%	5.85%	16.29%	6.63%	3.64%	8.78%
Highly skilled or skilled	53.89%	64.37%	73.89%	68.16%	47.66%	61.18%	66.49%	63.46%
Semi-skilled or unskilled	8.84%	15.44%	16.37%	6.46%	2.83%	22.98%	20.69%	14.34%
Apprentices	1.85%	4.84%	2.95%	12.72%	1.35%	3.67%	7.53%	3.10%
NUTS II								
North	24.22%	13.91%	11.98%	21.61%	18.72%	10.97%	4.74%	15.33%
Algarve	3.06%	7.60%	14.92%	7.35%	8.88%	6.10%	12.58%	9.60%
Centre	14.70%	14.08%	25.43%	11.31%	7.53%	8.87%	12.03%	19.22%
Lisbon	51.70%	59.25%	37.24%	47.35%	59.63%	70.53%	52.36%	40.58%
Alentejo	3.80%	4.37%	9.51%	10.05%	3.84%	2.68%	17.88%	11.40%
Azores	1.03%	0.27%	0.27%	0.97%	0.25%	0.38%	0.05%	1.09%
Madeira	1.50%	0.52%	0.66%	1.36%	1.15%	0.48%	0.35%	2.79%
CAE								
$\mathbf{CAE_1}$	1.08%	2.78%	9.36%	0.37%	1.42%	0.92%	19.84%	15.46%
$\mathbf{CAE_2}$	0.57%	0.17%	0.81%	0.03%	0.22%	0.29%	0.07%	0.25%
$\mathbf{CAE_3}$	26.49%	14.08%	21.63%	0.66%	15.28%	11.98%	6.43%	15.96%
$\mathbf{CAE_4}$	0.97%	0.07%	0.04%	0.00%	0.17%	0.10%	0.00%	0.13%
$\mathbf{CAE_5}$	1.78%	0.83%	2.29%	0.01%	0.35%	1.77%	0.68%	1.54%
$\mathbf{CAE_6}$	8.40%	15.41%	23.34%	0.28%	3.75%	27.55%	4.39%	11.76%
$\mathbf{CAE_7}$	13.18%	12.57%	9.34%	54.02%	9.19%	13.31%	12.63%	10.52%
$\mathbf{CAE_8}$	10.10%	7.53%	14.64%	0.12%	5.21%	4.42%	0.19%	5.76%
$\mathbf{CAE_9}$	5.19%	19.25%	5.57%	40.37%	6.78%	14.03%	41.06%	10.15%
$\mathbf{CAE_{10}}$	6.10%	4.83%	0.77%	0.37%	7.81%	1.99%	0.97%	3.59%
$\mathbf{CAE_{11}}$	6.03%	0.63%	0.11%	0.69%	3.40%	1.01%	0.13%	0.90%
$\mathbf{CAE_{12}}$	0.29%	0.82%	0.99%	0.74%	0.82%	0.66%	0.26%	0.58%
$\mathbf{CAE_{13}}$	3.09%	2.55%	1.29%	0.60%	7.24%	2.43%	0.51%	2.88%
$\mathbf{CAE_{14}}$	10.31%	12.18%	6.89%	0.46%	27.09%	14.36%	11.74%	13.56%
$\mathbf{CAE_{15}}$	0.26%	0.10%	0.09%	0.00%	0.08%	0.16%	0.00%	0.04%

CAE_{16}	0.71%	0.39%	0.37%	0.27%	3.67%	0.45%	0.08%	1.18%
CAE_{17}	3.67%	1.82%	0.86%	0.00%	3.92%	2.43%	0.15%	1.74%
CAE_{18}	0.96%	1.98%	0.97%	0.91%	2.14%	0.94%	0.30%	3.04%
CAE_{19}	0.82%	2.01%	0.64%	0.10%	1.45%	1.17%	0.56%	0.97%
CAE_{20}	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
CAE_{21}	0.00%	0.00%	0.00%	0.00%	0.01%	0.01%	0.00%	0.00%
FIRM SIZE								
Micro	2.68%	22.10%	23.14%	67.93%	10.76%	13.19%	23.12%	13.33%
Small	13.85%	33.48%	36.68%	29.67%	15.03%	28.81%	32.39%	25.48%
Medium	27.18%	19.73%	24.78%	1.42%	20.36%	24.66%	25.04%	28.65%
Large	56.30%	24.69%	15.40%	0.98%	53.85%	33.34%	19.45%	32.53%

Notes: Micro (less than 10 employees), small (10 to 49 employees), medium (50 to 249 employees) and large (more than 250 employees).

Source: Quadros de Pessoal. Own elaboration.

Table 16 - Mean of the real hourly wage for the female sample (2011 - 2020)

	Portugal	Brazil	CEEC	China	EU14	PALOP	SA	Others
Minimum wage earners (in %)	11.29%	37.17%	31.57%	60.96%	7.28%	28.75%	46.45%	24.20%
EDUCATION LEVELS (in €/h)								
Primary	3.77	3.38	3.33	3.34	4.19	3.39	3.42	3.40
Secondary	4.90	3.67	3.63	3.88	5.25	3.86	3.61	4.25
College	8.74	6.10	5.86	7.82	9.62	7.92	5.51	7.33
Mater	8.88	7.04	7.60	11.15	10.10	8.50	7.31	9.52
Doctorate	13.92	11.17	13.50	14.88	14.16	13.78	12.93	14.42
QUALIFICATION LEVELS (in €/h)								
Top or intermediary executives	10.53	8.57	9.21	7.20	12.90	10.02	8.82	10.71
Supervisors	7.22	5.28	5.71	5.19	8.10	5.75	4.97	7.23
Highly skilled or skilled	4.08	3.55	3.53	3.39	4.67	3.55	3.50	3.84
Semi-skilled or unskilled	3.31	3.33	3.25	3.16	3.47	3.24	3.38	3.29
Apprentices	3.36	3.25	3.24	3.14	3.69	3.27	3.38	3.31
NUTS II (in €/h)								
North	5.22	4.01	4.04	3.61	7.41	4.40	4.95	4.95
Algarve	4.80	3.60	3.60	3.27	6.53	3.66	3.64	3.83
Centre	4.50	3.64	3.45	3.44	6.22	4.07	3.58	4.17
Lisbon	6.36	3.88	4.20	3.93	7.42	3.85	3.64	5.65
Alentejo	4.75	3.54	3.57	3.50	5.94	3.99	3.42	3.59
Azores	5.10	3.93	3.70	3.40	8.53	3.73	4.21	4.51
Madeira	5.43	4.27	4.64	3.17	7.84	4.15	3.45	4.82
CAE (in €/h)								
CAE_1	3.92	3.42	3.32	5.94	5.66	3.60	3.44	3.28
CAE_2	7.57	4.75	3.49	-	13.52	4.97	-	16.41
CAE_3	4.81	3.79	3.56	6.67	7.17	3.80	3.51	4.60
CAE_4	13.17	18.81	11.40	-	21.10	22.94	-	9.74
CAE_5	6.47	3.86	3.47	-	7.02	3.44	3.20	3.76
CAE_6	6.32	4.39	3.78	5.24	6.53	3.94	3.27	4.43
CAE_7	4.85	3.69	3.71	3.49	8.30	3.90	3.47	4.73
CAE_8	7.74	4.16	4.42	6.47	9.92	7.58	7.22	5.47
CAE_9	4.17	3.50	3.57	3.33	5.99	3.44	3.49	3.77
CAE_{10}	9.10	7.35	7.59	7.00	9.09	8.23	9.96	8.01
CAE_{11}	9.17	7.09	7.24	13.63	11.01	8.41	7.18	9.83
CAE_{12}	6.35	3.85	3.66	5.29	6.41	4.23	3.44	4.25
CAE_{13}	7.76	5.23	5.27	6.67	8.01	5.32	6.96	7.16
CAE_{14}	4.35	3.55	3.71	5.43	4.71	3.31	3.64	4.17
CAE_{15}	8.28	3.82	3.43	-	11.36	3.94	-	7.43
CAE_{16}	7.13	5.18	6.69	6.99	11.10	4.38	8.22	10.33
CAE_{17}	5.63	3.84	4.44	4.62	8.54	3.99	3.43	5.16
CAE_{18}	5.91	4.59	4.50	6.96	7.78	4.48	3.49	5.72

CAE_{19}	5.84	3.51	3.74	6.36	8.89	3.98	3.84	5.33
CAE_{20}	-	-	-	-	-	-	-	-
CAE_{21}	12.93	-	5.91	-	12.73	14.93	8.24	25.83
FIRM SIZE (in €/h)								
Micro	4.08	3.47	3.39	3.42	5.85	3.53	3.43	4.09
Small	4.77	3.68	3.57	3.87	7.20	3.75	3.56	4.20
Medium	5.35	4.02	3.97	11.58	8.82	3.90	3.87	5.40
Large	6.09	4.19	4.71	7.87	6.87	4.13	3.84	5.35

Notes: Micro (less than 10 employees), small (10 to 49 employees), medium (50 to 249 employees) and large (more than 250 employees).

Source: Quadros de Pessoal. Own elaboration.

Table 17 - Mean of the real hourly wage for the male sample (2011 - 2020)

	Portugal	Brazil	CEEC	China	EU14	PALOP	SA	Others
Minimum wage earners (in %)	6.31%	28.31%	22.09%	62.49%	5.54%	22.30%	47.13%	21.94%
EDUCATION LEVELS (in €/h)								
Primary	4.74	4.53	3.80	3.34	7.58	3.77	3.39	5.40
Secondary	6.27	5.81	4.62	3.65	12.84	4.44	3.62	12.10
College	11.81	9.93	6.83	19.78	16.12	9.93	5.68	11.86
Mater	10.84	10.90	9.79	29.68	15.22	10.14	9.12	12.29
Doctorate	16.31	14.88	17.39	17.47	19.69	9.36	14.32	16.13
QUALIFICATION LEVELS (in €/h)								
Top or intermediary executives	13.27	12.14	10.52	12.33	20.11	11.87	9.80	16.02
Supervisors	8.09	18.28	11.29	4.55	26.16	7.40	4.74	42.17
Highly skilled or skilled	4.64	3.91	3.85	3.31	5.17	3.79	3.45	4.01
Semi-skilled or unskilled	3.67	3.48	3.41	3.16	4.03	3.40	3.37	3.40
Apprentices	3.60	3.40	3.38	3.12	3.65	3.35	3.34	3.43
NUTS II (in €/h)								
North	6.18	7.95	4.33	3.52	15.72	4.99	3.98	13.30
Algarve	5.17	4.17	3.75	3.31	8.47	3.83	3.58	4.46
Centre	5.37	4.22	3.81	3.58	9.74	4.07	3.53	4.50
Lisbon	7.49	5.52	4.86	4.51	13.98	4.34	3.60	10.92
Alentejo	5.75	4.07	3.76	3.40	8.78	4.48	3.41	4.06
Azores	6.07	8.35	3.68	3.44	14.58	5.08	4.16	5.56
Madeira	6.01	12.50	4.86	3.36	14.08	6.15	4.13	7.58
CAE (in €/h)								
CAE_1	4.69	3.70	3.52	3.60	8.26	4.26	3.41	3.38
CAE_2	6.36	5.54	4.28	5.27	22.24	4.45	6.10	24.58
CAE_3	6.23	4.75	4.07	6.66	12.73	4.75	3.67	5.87
CAE_4	12.58	10.84	5.59	-	13.18	9.38	-	10.43
CAE_5	5.06	3.93	3.77	12.59	7.79	3.71	3.30	4.24
CAE_6	5.49	4.12	3.65	4.66	8.27	3.74	3.51	4.17
CAE_7	6.14	4.38	4.20	3.75	14.40	4.53	3.46	5.97
CAE_8	6.47	4.05	3.75	3.58	15.99	5.01	5.01	4.84
CAE_9	5.00	3.73	3.92	3.26	8.78	3.75	3.55	4.27
CAE_{10}	10.22	9.42	9.29	8.75	10.92	7.97	6.57	9.76
CAE_{11}	10.70	9.50	10.49	35.07	21.37	10.33	11.11	13.12
CAE_{12}	7.85	4.23	4.00	5.43	10.79	4.38	3.72	4.94
CAE_{13}	9.25	8.15	6.10	15.24	13.07	5.67	7.04	11.28
CAE_{14}	4.56	3.85	3.78	8.00	5.18	3.61	3.43	4.15
CAE_{15}	7.45	5.14	4.85	-	7.57	4.13	-	6.72
CAE_{16}	8.41	7.75	9.54	15.68	12.91	5.16	13.41	12.81
CAE_{17}	7.23	5.28	7.13	-	13.51	5.93	3.53	6.95
CAE_{18}	10.85	54.44	32.84	5.74	133.44	11.84	4.04	109.76

CAE_{19}	7.35	4.18	4.61	6.50	13.46	4.59	3.81	7.21
CAE_{20}	-	-	-	-	-	-	-	-
CAE_{21}	11.37	26.44	6.99	-	12.98	28.65	-	-
FIRM SIZE (in €/h)								
Micro	4.48	3.80	3.63	3.43	7.69	3.82	3.33	4.44
Small	5.42	4.25	3.89	4.10	10.51	3.98	3.56	4.79
Medium	6.56	8.84	4.83	20.23	22.40	4.63	3.62	13.51
Large	7.17	6.24	5.11	13.25	11.76	4.73	3.81	8.65

Notes: Micro (less than 10 employees), small (10 to 49 employees), medium (50 to 249 employees) and large (more than 250 employees).

Source: Quadros de Pessoal. Own elaboration.

Annex C - Regression diagnostic

Table 18 - Statistical tests

Regression Diagnostic	Test	Specification	Chi-square	P-value
Heteroscedasticity	Breusch–Pagan	1	958,797.91	0.0000
		2	956,804.66	0.0000
		3	1.05E+06	0.0000
		4	1.64E+06	0.0000
		5	1.67E+06	0.0000
Specification	Hausman	4	106,837.07	0.0000

Source: Quadros de Pessoal. Own elaboration.

Annex D - Results of the OLS and FE estimations

Table 19 - Pooled OLS wage regressions – female sample (2011 - 2020)

(Dependent variable: log of the real hourly wage)

	Specification 1	Specification 2	Specification 3	Specification 4	Specification 5
Immigrant	-0.0731*	-0.0797*	-0.00933*	0.0410*	0.161*
	(0.0011)	(0.00110)	(0.00110)	(0.000889)	(0.00199)
YSM	-0.0092*	-0.00824*	-0.00885*	-0.00973*	-0.00745*
	(0.0003)	(0.000263)	(0.000252)	(0.000210)	(0.000214)
YSM^2	0.0004*	0.000410*	0.000394*	0.000307*	0.000189*
	(0.00001)	(1.22e-05)	(1.16e-05)	(9.94e-06)	(9.98e-06)
Age	0.0363*	0.0366*	0.0303*	0.0177*	0.0179*
	(0.00011)	(0.000113)	(0.000108)	(9.38e-05)	(9.38e-05)
Age^2	-0.0003*	-0.000285*	-0.000229*	-0.000155*	-0.000158*
	(1.42e-06)	(1.42e-06)	(1.35e-06)	(1.16e-06)	(1.16e-06)
$Educ_2$	0.2645*	0.264*	0.220*	0.146*	0.145*
	(0.00040)	(0.000409)	(0.000391)	(0.000318)	(0.000319)
$Educ_3$	0.7976*	0.798*	0.715*	0.386*	0.383*
	(0.00051)	(0.000516)	(0.000589)	(0.000568)	(0.000568)
$Educ_4$	0.8842*	0.882*	0.778*	0.424*	0.421*
	(0.00134)	(0.00135)	(0.00139)	(0.00126)	(0.00126)
$Educ_5$	1.2618*	1.257*	1.197*	0.746*	0.742*
	(0.00348)	(0.00348)	(0.00355)	(0.00327)	(0.00328)
Algarve	0.0084*	0.00920*	0.0728*	0.0643*	0.0647*
	(0.00080)	(0.000804)	(0.000826)	(0.000684)	(0.000684)
Centre	-0.0695*	-0.0700*	-0.0436*	-0.0319*	-0.0316*
	(0.00049)	(0.000493)	(0.000498)	(0.000416)	(0.000417)
Lisbon	0.0910*	0.0916*	0.0917*	0.0751*	0.0764*
	(0.00042)	(0.000422)	(0.000454)	(0.000370)	(0.000370)
Alentejo	-0.0353*	-0.0349*	0.00112	0.00983*	0.0101*
	(0.00081)	(0.000811)	(0.000810)	(0.000665)	(0.000666)
Azores	0.0624*	0.0633*	0.0484*	0.0265*	0.0252*
	(0.00150)	(0.0015)	(0.00144)	(0.00140)	(0.00139)
Madeira	0.0830*	0.0822*	0.106*	0.0727*	0.0717*
	(0.00127)	(0.00127)	(0.00126)	(0.00103)	(0.00103)

$Year_{12}$	-	-0.0357* (0.00083)	-0.0341* (0.000779)	-0.0340* (0.000644)	-0.0341* (0.000643)
$Year_{13}$	-	-0.0487* (0.000841)	-0.0454* (0.000791)	-0.0432* (0.000651)	-0.0435* (0.000650)
$Year_{14}$	-	-0.0515* (0.000836)	-0.0453* (0.000787)	-0.0406* (0.000646)	-0.0409* (0.000645)
$Year_{15}$	-	-0.0596* (0.000828)	-0.0530* (0.000781)	-0.0461* (0.000640)	-0.0464* (0.000639)
$Year_{16}$	-	-0.0588* (0.000814)	-0.0495* (0.000771)	-0.0417* (0.000634)	-0.0422* (0.000633)
$Year_{17}$	-	-0.0565* (0.000796)	-0.0436* (0.000754)	-0.0332* (0.000621)	-0.0338* (0.000621)
$Year_{18}$	-	-0.0433* (0.00078)	-0.0294* (0.000740)	-0.0166* (0.000611)	-0.0169* (0.000610)
$Year_{19}$	-	-0.0283* (0.000765)	-0.00982* (0.000726)	0.00853* (0.000605)	0.00850* (0.000605)
$Year_{20}$	-	-0.00478* (0.000761)	0.0191* (0.000724)	0.0343* (0.000607)	0.0344* (0.000607)
Concentration	-	-	-0.0203* (0.000188)	-0.00811* (0.000156)	-0.00881* (0.000160)
Size	-	-	0.00769* (0.000111)	0.00142* (9.17e-05)	0.000749* (9.25e-05)
CAE_1	-	-	-0.0659* (0.00130)	-0.0690* (0.00118)	-0.0640* (0.00121)
CAE_2	-	-	0.254* (0.00757)	0.197* (0.00658)	0.197* (0.00657)
CAE_4	-	-	0.604* (0.00396)	0.452* (0.00331)	0.452* (0.00331)
CAE_5	-	-	0.0223* (0.00241)	0.0137* (0.00196)	0.0151* (0.00196)
CAE_6	-	-	0.0440* (0.00184)	-0.00775* (0.00165)	-0.00660* (0.00165)
CAE_7	-	-	-0.0438* (0.000591)	-0.0261* (0.000490)	-0.0253* (0.000492)
CAE_8	-	-	0.183* (0.00128)	0.138* (0.00107)	0.139* (0.00107)
CAE_9	-	-	-0.0415* (0.000648)	-0.0251* (0.000538)	-0.0220* (0.000539)

CAE_{10}	-	-	0.193* (0.00131)	0.112* (0.00113)	0.112* (0.00113)
CAE_{11}	-	-	0.245* (0.000923)	0.234* (0.000766)	0.235* (0.000768)
CAE_{12}	-	-	0.0324* (0.00298)	0.0352* (0.00264)	0.0339* (0.00264)
CAE_{13}	-	-	0.0746* (0.00134)	0.0413* (0.00115)	0.0401* (0.00115)
CAE_{14}	-	-	-0.164* (0.000752)	-0.0690* (0.000634)	-0.0716* (0.000635)
CAE_{15}	-	-	0.173* (0.00453)	0.113* (0.00356)	0.112* (0.00356)
CAE_{16}	-	-	0.0295* (0.00125)	-0.0118* (0.00102)	-0.0138* (0.00102)
CAE_{17}	-	-	-0.0326* (0.000516)	-0.0633* (0.000439)	-0.0624* (0.000439)
CAE_{18}	-	-	0.0196* (0.00234)	0.0346* (0.00192)	0.0347* (0.00191)
CAE_{19}	-	-	-0.0135* (0.00117)	-0.0262* (0.000975)	-0.0254* (0.000974)
CAE_{21}	-	-	0.654* (0.0231)	0.620* (0.0229)	0.618* (0.0229)
Tenure	-	-	-	0.0104* (5.24e-05)	0.0103* (5.25e-05)
$Tenure^2$	-	-	-	-4.79e-05* (1.76e-06)	-4.58e-05* (1.76e-06)
$Qual_1$	-	-	-	0.618* (0.000885)	0.616* (0.000883)
$Qual_2$	-	-	-	0.339* (0.000761)	0.338* (0.000759)
$Qual_3$	-	-	-	0.0354* (0.000579)	0.0343* (0.000578)
$Qual_4$	-	-	-	-0.0147* (0.000615)	-0.0135* (0.000615)
Brazil	-	-	-	-	-0.165* (0.00245)
CEEC	-	-	-	-	-0.163* (0.00269)

China	-	-	-	-	-0.151* (0.00389)
PALOP	-	-	-	-	-0.159* (0.00287)
SA	-	-	-	-	-0.0981* (0.00432)
Others	-	-	-	-	-0.137* (0.00424)
Brazil · CIMMIG	-	-	-	-	-0.00266* (0.000610)
CEEC · CIMMIG	-	-	-	-	-6.40e-05 (0.000807)
China · CIMMIG	-	-	-	-	-0.0485* (0.00461)
PALOP · CIMMIG	-	-	-	-	-0.00825* (0.000745)
SA · CIMMIG	-	-	-	-	0.000456 (0.00250)
Others · CIMMIG	-	-	-	-	-0.00712* (0.00157)
Constant	0.248* (0.00220)	0.281* (0.00225)	0.329* (0.00228)	0.643* (0.00196)	0.643* (0.00197)
N	4.078.776	4,078,776	4,078,776	4,078,776	4,078,776
R^2	0.5043	0.506	0.551	0.695	0.696

Notes: (i) In parentheses are the robust standard errors.

(ii) *, **, *** denote significant at 1%, 5%, and 10%, respectively.

(iii) CIMMIG is the "Concentration" variable.

Source: Quadros de Pessoal. Own elaboration.

Table 20 - Pooled OLS wage regressions – male sample (2011 - 2020)

(Dependent variable: log of the real hourly wage)

	Specification 1	Specification 2	Specification 3	Specification 4	Specification 5
Immigrant	-0.0977*	-0.0981*	0.000582	0.0467*	0.235*
	(0.00127)	(0.00128)	(0.00125)	(0.00109)	(0.00277)
YSM	-0.0155*	-0.0156*	-0.0162*	-0.0139*	-0.0104*
	(0.00026)	(0.000267)	(0.000253)	(0.000214)	(0.000223)
YSM^2	0.0006*	0.000599*	0.000581*	0.000424*	0.000290*
	(0.0000)	(1.17e-05)	(1.12e-05)	(9.31e-06)	(9.24e-06)
Age	0.0429*	0.0427*	0.0401*	0.0225*	0.0225*
	(0.0001)	(0.000124)	(0.000117)	(0.000110)	(0.000110)
Age^2	-0.0003*	-0.000317*	-0.000306*	-0.000194*	-0.000196*
	(1.54e-06)	(1.55e-06)	(1.46e-06)	(1.35e-06)	(1.35e-06)
$Educ_2$	0.2755*	0.278*	0.222*	0.135*	0.133*
	(0.0004)	(0.000453)	(0.000444)	(0.000370)	(0.000370)
$Educ_3$	0.8393*	0.842*	0.747*	0.401*	0.396*
	(0.0006)	(0.000645)	(0.000739)	(0.000726)	(0.000726)
$Educ_4$	0.9004*	0.906*	0.785*	0.399*	0.394*
	(0.0014)	(0.00140)	(0.00148)	(0.00143)	(0.00143)
$Educ_5$	1.2188*	1.220*	1.171*	0.702*	0.689*
	(0.0039)	(0.00388)	(0.00407)	(0.00388)	(0.00385)
Algarve	-0.0369*	-0.0357*	0.0678*	0.0645*	0.0644*
	(0.0010)	(0.000991)	(0.00101)	(0.000873)	(0.000874)
Centre	-0.0475*	-0.0473*	-0.0184*	-0.00541*	-0.00447*
	(0.0006)	(0.000574)	(0.000566)	(0.000491)	(0.000490)
Lisbon	0.0803*	0.0801*	0.117*	0.0950*	0.0972*
	(0.0005)	(0.000476)	(0.000503)	(0.000429)	(0.000429)
Alentejo	0.0029**	0.00450*	0.0486*	0.0536*	0.0559*
	(0.0009)	(0.000959)	(0.000944)	(0.000798)	(0.000798)
Azores	0.0381*	0.0367*	-0.00466**	-0.00880*	-0.0108*
	(0.0019)	(0.00189)	(0.00184)	(0.00166)	(0.00166)
Madeira	0.0559*	0.0540*	0.0812*	0.0682*	0.0658*
	(0.0016)	(0.00155)	(0.00153)	(0.00131)	(0.00130)
$Year_{12}$	-	-0.0340*	-0.0340*	-0.0341*	-0.0342*
		(0.000892)	(0.000845)	(0.000721)	(0.000719)
$Year_{13}$	-	-0.0516*	-0.0517*	-0.0479*	-0.0481*
		(0.000904)	(0.000855)	(0.000729)	(0.000727)

$Year_{14}$	-	-0.0614*	-0.0590*	-0.0482*	-0.0485*
		(0.000898)	(0.000852)	(0.000725)	(0.000723)
$Year_{15}$	-	-0.0714*	-0.0703*	-0.0563*	-0.0565*
		(0.000894)	(0.000848)	(0.000721)	(0.000719)
$Year_{16}$	-	-0.0798*	-0.0754*	-0.0581*	-0.0585*
		(0.000880)	(0.000837)	(0.000711)	(0.000710)
$Year_{17}$	-	-0.0875*	-0.0777*	-0.0560*	-0.0564*
		(0.000863)	(0.000820)	(0.000698)	(0.000697)
$Year_{18}$	-	-0.0764*	-0.0645*	-0.0385*	-0.0386*
		(0.000844)	(0.000805)	(0.000687)	(0.000685)
$Year_{19}$	-	-0.0605*	-0.0429*	-0.0116*	-0.0115*
		(0.000831)	(0.000794)	(0.000683)	(0.000681)
$Year_{20}$	-	-0.0413*	-0.0158*	0.0127*	0.0130*
		(0.000818)	(0.000783)	(0.000677)	(0.000676)
Concentration	-	-	-0.0303*	-0.0125*	-0.0130*
			(0.000193)	(0.000167)	(0.000171)
Size	-	-	0.0112*	0.00535*	0.00417*
			(0.000133)	(0.000112)	(0.000112)
CAE_{1}	-	-	-0.131*	-0.114*	-0.104*
			(0.00128)	(0.00115)	(0.00118)
CAE_{2}	-	-	0.115*	0.134*	0.134*
			(0.00212)	(0.00185)	(0.00184)
CAE_{4}	-	-	0.371*	0.292*	0.294*
			(0.00145)	(0.00133)	(0.00133)
CAE_{5}	-	-	-0.213*	-0.135*	-0.134*
			(0.00120)	(0.000982)	(0.000980)
CAE_{6}	-	-	-0.110*	-0.0887*	-0.0852*
			(0.000666)	(0.000580)	(0.000581)
CAE_{7}	-	-	-0.0887*	-0.0688*	-0.0673*
			(0.000646)	(0.000540)	(0.000540)
CAE_{8}	-	-	-0.0951*	-0.0831*	-0.0815*
			(0.000717)	(0.000611)	(0.000611)
CAE_{9}	-	-	-0.0905*	-0.0895*	-0.0853*
			(0.000813)	(0.000684)	(0.000688)
CAE_{10}	-	-	0.0940*	-0.0193*	-0.0182*
			(0.00102)	(0.000947)	(0.000945)
CAE_{11}	-	-	0.138*	0.0622*	0.0658*
			(0.000960)	(0.000815)	(0.000814)

CAE$_{12}$	-	-	-0.0159*	-0.00562***	-0.00567***
			(0.00385)	(0.00341)	(0.00342)
CAE$_{13}$	-	-	0.00712*	-0.0437*	-0.0454*
			(0.00134)	(0.00119)	(0.00119)
CAE$_{14}$	-	-	-0.281*	-0.135*	-0.139*
			(0.000688)	(0.000610)	(0.000613)
CAE$_{15}$	-	-	-0.0491*	-0.0859*	-0.0864*
			(0.00350)	(0.00283)	(0.00284)
CAE$_{16}$	-	-	-0.0952*	-0.104*	-0.110*
			(0.00219)	(0.00195)	(0.00194)
CAE$_{17}$	-	-	-0.157*	-0.183*	-0.181*
			(0.00105)	(0.000887)	(0.000885)
CAE$_{18}$	-	-	0.249*	0.235*	0.234*
			(0.00431)	(0.00385)	(0.00383)
CAE$_{19}$	-	-	-0.0825*	-0.0876*	-0.0882*
			(0.00223)	(0.00200)	(0.00199)
CAE$_{21}$	-	-	0.482*	0.484*	0.481*
			(0.0198)	(0.0162)	(0.0171)
Tenure	-	-	-	0.0157*	0.0157*
				(5.62e-05)	(5.62e-05)
Tenure^2	-	-	-	-0.000190*	-0.000186*
				(1.78e-06)	(1.78e-06)
Qual$_1$	-	-	-	0.696*	0.692*
				(0.00100)	(0.00100)
Qual$_2$	-	-	-	0.379*	0.375*
				(0.000845)	(0.000841)
Qual$_3$	-	-	-	0.0694*	0.0667*
				(0.000651)	(0.000651)
Qual$_4$	-	-	-	-0.0204*	-0.0185*
				(0.000709)	(0.000711)
Brazil	-	-	-	-	-0.230*
					(0.00336)
CEEC	-	-	-	-	-0.251*
					(0.00317)
China	-	-	-	-	-0.316*
					(0.00451)
PALOP	-	-	-	-	-0.268*
					(0.00352)

SA	-	-	-	-	-0.241* (0.00302)
Others	-	-	-	-	-0.221* (0.00468)
Brazil · CIMMIG	-	-	-	-	-0.00783* (0.000792)
CEEC · CIMMIG	-	-	-	-	0.00107 (0.000766)
China · CIMMIG	-	-	-	-	-0.0795* (0.00613)
PALOP · CIMMIG	-	-	-	-	-0.00891* (0.000844)
SA · CIMMIG	-	-	-	-	-0.00473* (0.000967)
Others · CIMMIG	-	-	-	-	-0.0156* (0.00143)
Constant	0.238* (0.00238)	0.296* (0.00245)	0.253* (0.00243)	0.619* (0.00222)	0.627* (0.00223)
N	4.791.324	4,791,324	4,791,324	4,791,324	4,791,324
R^2	0.4392	0.441	0.493	0.630	0.632

Notes: (i) In parentheses are the robust standard errors.

(ii) *, **, *** denote significant at 1%, 5%, and 10%, respectively.

(iii) CIMMIG is the "Concentration" variable.

Source: Quadros de Pessoal. Own elaboration.

Table 21 - Unit fixed effects wage regressions (2011 - 2020)

(Dependent variable: log of the real hourly wage)

	Worker FE		Firm FE		Qualification level FE	
	Women	**Men**	**Women**	**Men**	**Women**	**Men**
Immigrant	0.0410*	0.0467*	-0.0155*	-0.0139*	-0.0152*	-0.0134*
	(0.000910)	(0.000932)	(0.00104)	(0.00111)	(0.00103)	(0.00110)
YSM	-0.00973*	-0.0139*	0.00222*	0.00153*	0.00224*	0.00150*
	(0.000198)	(0.000190)	(0.000154)	(0.000160)	(0.000154)	(0.000159)
YSM^2	0.000307*	0.000424*	-5.63e-05*	-2.97e-05*	-5.75e-05*	-2.91e-05*
	(8.51e-06)	(7.78e-06)	(6.18e-06)	(6.12e-06)	(6.15e-06)	(6.09e-06)
Age	0.0177*	0.0225*	0.0334*	0.0519*	0.0328*	0.0508*
	(0.000102)	(0.000111)	(0.000101)	(0.000112)	(0.000101)	(0.000112)
Age^2	-0.000155*	-0.000194*	-0.000194*	-0.000405*	-0.000188*	-0.000394*
	(1.23e-06)	(1.33e-06)	(1.19e-06)	(1.31e-06)	(1.19e-06)	(1.31e-06)
$Educ_2$	0.146*	0.135*	0.000606	0.00420*	0.000357	0.00411*
	(0.000361)	(0.000391)	(0.000419)	(0.000424)	(0.000417)	(0.000423)
$Educ_3$	0.386*	0.401*	0.0307*	0.0395*	0.0284*	0.0360*
	(0.000481)	(0.000580)	(0.000665)	(0.000753)	(0.000662)	(0.000750)
$Educ_4$	0.424*	0.399*	0.0455*	0.0687*	0.0426*	0.0637*
	(0.000948)	(0.00113)	(0.000944)	(0.00111)	(0.000940)	(0.00110)
$Educ_5$	0.746*	0.702*	0.0628*	0.0799*	0.0595*	0.0741*
	(0.00235)	(0.00301)	(0.00220)	(0.00277)	(0.00219)	(0.00275)
Algarve	0.0643*	0.0645*	0.0224*	0.00504	0.0238*	0.00715***
	(0.000753)	(0.000918)	(0.00379)	(0.00425)	(0.00378)	(0.00423)
Centre	-0.0319*	-0.00541*	-0.00902*	-0.0110*	-0.00792*	-0.0102*
	(0.000474)	(0.000528)	(0.00190)	(0.00182)	(0.00189)	(0.00181)
Lisbon	0.0751*	0.0950*	0.00389*	0.00422*	0.00522*	0.00472*
	(0.000369)	(0.000421)	(0.00136)	(0.00136)	(0.00135)	(0.00136)
Alentejo	0.00983*	0.0536*	0.00382	0.00424***	0.00654**	0.00726*
	(0.000772)	(0.000864)	(0.00282)	(0.00234)	(0.00281)	(0.00233)
Azores	0.0265*	-0.00880*	-0.0200*	-0.0252*	-0.0345*	-0.0222*
	(0.00141)	(0.00162)	(0.00393)	(0.00853)	(0.00392)	(0.00849)
Madeira	0.0727*	0.0682*	0.0260*	-0.0138*	0.0276*	-0.0126*
	(0.00115)	(0.00134)	(0.00370)	(0.00423)	(0.00368)	(0.00421)
Concentration	-0.00811*	-0.0125*	0.0129*	0.0120*	0.0128*	0.0118*
	(0.000147)	(0.000162)	(0.000124)	(0.000139)	(0.000123)	(0.000138)

Size	0.00142*	0.00535*	0.0485*	0.0605*	0.0477*	0.0597*
	(8.88e-05)	(0.000110)	(0.000272)	(0.000284)	(0.000271)	(0.000283)
CAE_1	-0.0690*	-0.114*	-0.00454	0.0115*	-0.00559***	0.0109*
	(0.00136)	(0.00132)	(0.00294)	(0.00334)	(0.00293)	(0.00332)
CAE_2	0.197*	0.134*	-0.0375**	-0.0728*	-0.0371**	-0.0690*
	(0.00507)	(0.00215)	(0.0175)	(0.00918)	(0.0174)	(0.00913)
CAE_4	0.452*	0.292*	-0.0816*	-0.00694	-0.0752*	-0.00343
	(0.00319)	(0.00174)	(0.0181)	(0.0177)	(0.0180)	(0.0176)
CAE_5	0.0137*	-0.135*	0.00753	-0.0115***	0.00901	-0.0133**
	(0.00189)	(0.00123)	(0.0103)	(0.00633)	(0.0102)	(0.00630)
CAE_6	-0.00775*	-0.0887*	0.00788**	0.0141*	0.00835*	0.0141*
	(0.00131)	(0.000627)	(0.00311)	(0.00184)	(0.00309)	(0.00183)
CAE_7	-0.0261*	-0.0688*	0.0111*	0.0129*	0.0120*	0.0136*
	(0.000483)	(0.000545)	(0.00139)	(0.00141)	(0.00139)	(0.00140)
CAE_8	0.138*	-0.0831*	0.00354	-0.000939	0.00428	-0.00110
	(0.000867)	(0.000606)	(0.00446)	(0.00301)	(0.00444)	(0.00300)
CAE_9	-0.0251*	-0.0895*	-0.00704*	0.00372	-0.00730*	0.00366
	(0.000601)	(0.000759)	(0.00194)	(0.00251)	(0.00193)	(0.00249)
CAE_{10}	0.112*	-0.0193*	0.0458*	0.0319*	0.0473*	0.0315*
	(0.000846)	(0.000780)	(0.00284)	(0.00249)	(0.00283)	(0.00248)
CAE_{11}	0.234*	0.0622*	-0.0239*	-0.00862	-0.0208*	-0.00583
	(0.000696)	(0.000795)	(0.00645)	(0.00752)	(0.00642)	(0.00748)
CAE_{12}	0.0352*	-0.00562**	-0.0236*	0.00900***	-0.0252*	0.00692
	(0.00207)	(0.00275)	(0.00357)	(0.00472)	(0.00355)	(0.00469)
CAE_{13}	0.0413*	-0.0437*	-0.0121*	-0.0281*	-0.00910*	-0.0273*
	(0.000874)	(0.000983)	(0.00237)	(0.00226)	(0.00235)	(0.00225)
CAE_{14}	-0.0690*	-0.135*	0.0156*	0.00413**	0.0163*	0.00430**
	(0.000633)	(0.000666)	(0.00221)	(0.00201)	(0.00220)	(0.00200)
CAE_{15}	0.113*	-0.0859*	0.0662**	0.0560	0.0674**	0.0565***
	(0.00292)	(0.00317)	(0.0284)	(0.0344)	(0.0283)	(0.0342)
CAE_{16}	-0.0118*	-0.104*	0.0148*	0.00760	0.0154*	0.00899
	(0.00100)	(0.00191)	(0.00230)	(0.00571)	(0.00229)	(0.00569)
CAE_{17}	-0.0633*	-0.183*	0.0171*	0.0280*	0.0173*	0.0286*
	(0.000460)	(0.000913)	(0.00212)	(0.00513)	(0.00211)	(0.00511)
CAE_{18}	0.0346*	0.235*	0.0136*	0.0230*	0.0148*	0.0223*
	(0.00162)	(0.00159)	(0.00426)	(0.00490)	(0.00424)	(0.00487)
CAE_{19}	-0.0262*	-0.0876*	0.0202*	0.0442*	0.0201*	0.0448*
	(0.000935)	(0.00172)	(0.00176)	(0.00283)	(0.00176)	(0.00282)

CAE_{21}	0.620* (0.0200)	0.484* (0.0277)	-	-	-	-
Tenure	0.0104* (5.01e-05)	0.0157* (5.49e-05)	0.00730* (4.49e-05)	0.00563* (4.71e-05)	0.00708* (4.47e-05)	0.00553* (4.69e-05)
$Tenure^2$	-4.79e-05* (1.60e-06)	-0.000190* (1.70e-06)	-0.000272* (1.43e-06)	-0.000186* (1.51e-06)	-0.000267* (1.42e-06)	-0.000186* (1.50e-06)
$Qual_1$	0.618* (0.00104)	0.696* (0.00124)	0.143* (0.000675)	0.157* (0.000764)	-	-
$Qual_2$	0.339* (0.00101)	0.379* (0.00119)	0.102* (0.000613)	0.109* (0.000699)	-	-
$Qual_3$	0.0354* (0.000936)	0.0694* (0.00112)	0.0367* (0.000558)	0.0435* (0.000644)	-	-
$Qual_4$	-0.0147* (0.00103)	-0.0204* (0.00123)	0.00925* (0.000645)	0.0128* (0.000747)	-	-
Constant	0.624* (0.00220)	0.587* (0.00241)	0.233* (0.00297)	-0.0978* (0.00316)	0.316* (0.00293)	0.00256 (0.00313)
N	4,078,776	4,791,324	3,797,049	4,452,582	3,797,049	4,452,582
R-squared	0.693	0.629	0.978	0.973	0.978	0.973

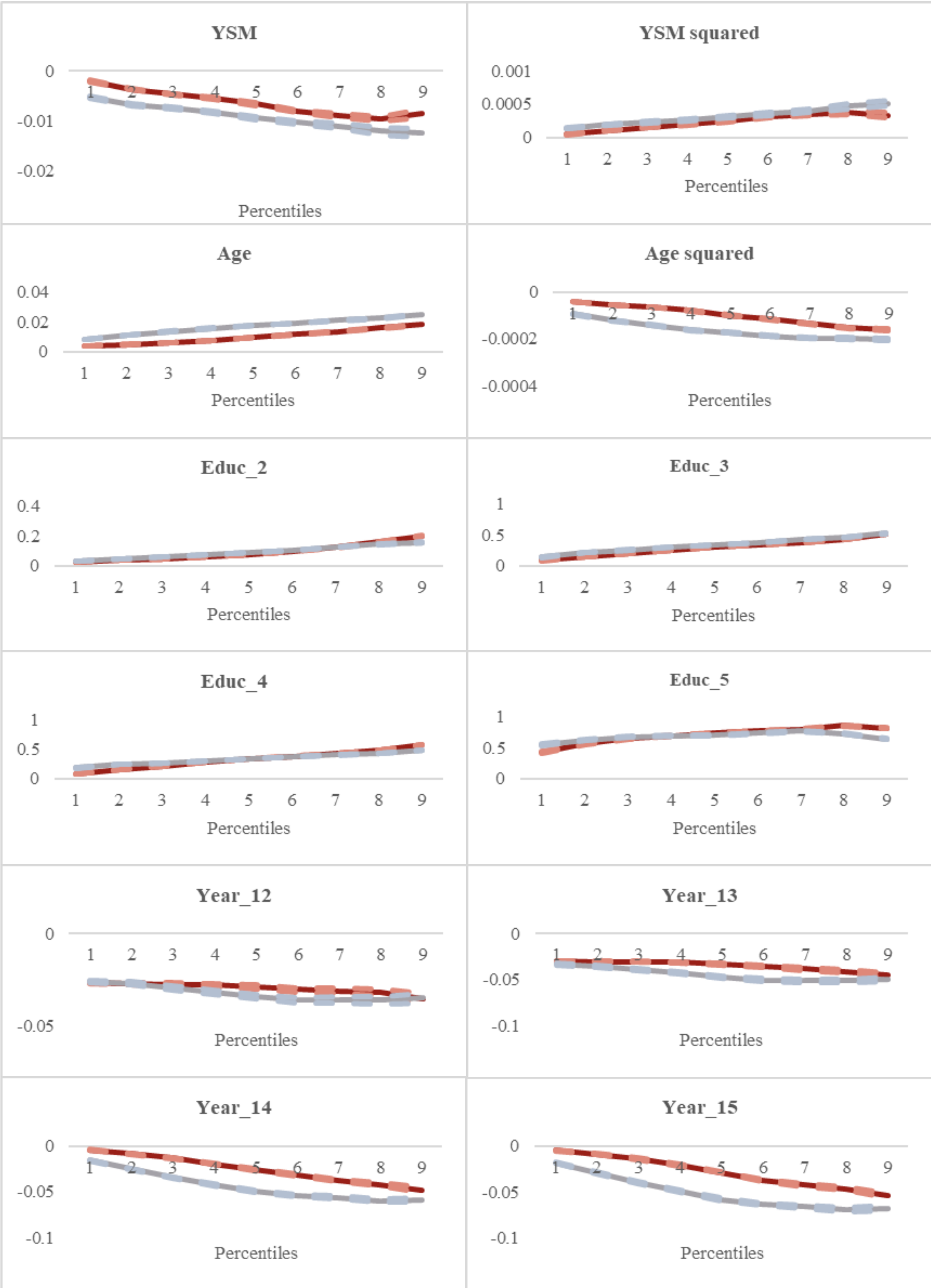
Notes: (i) In parentheses are the robust standard errors.

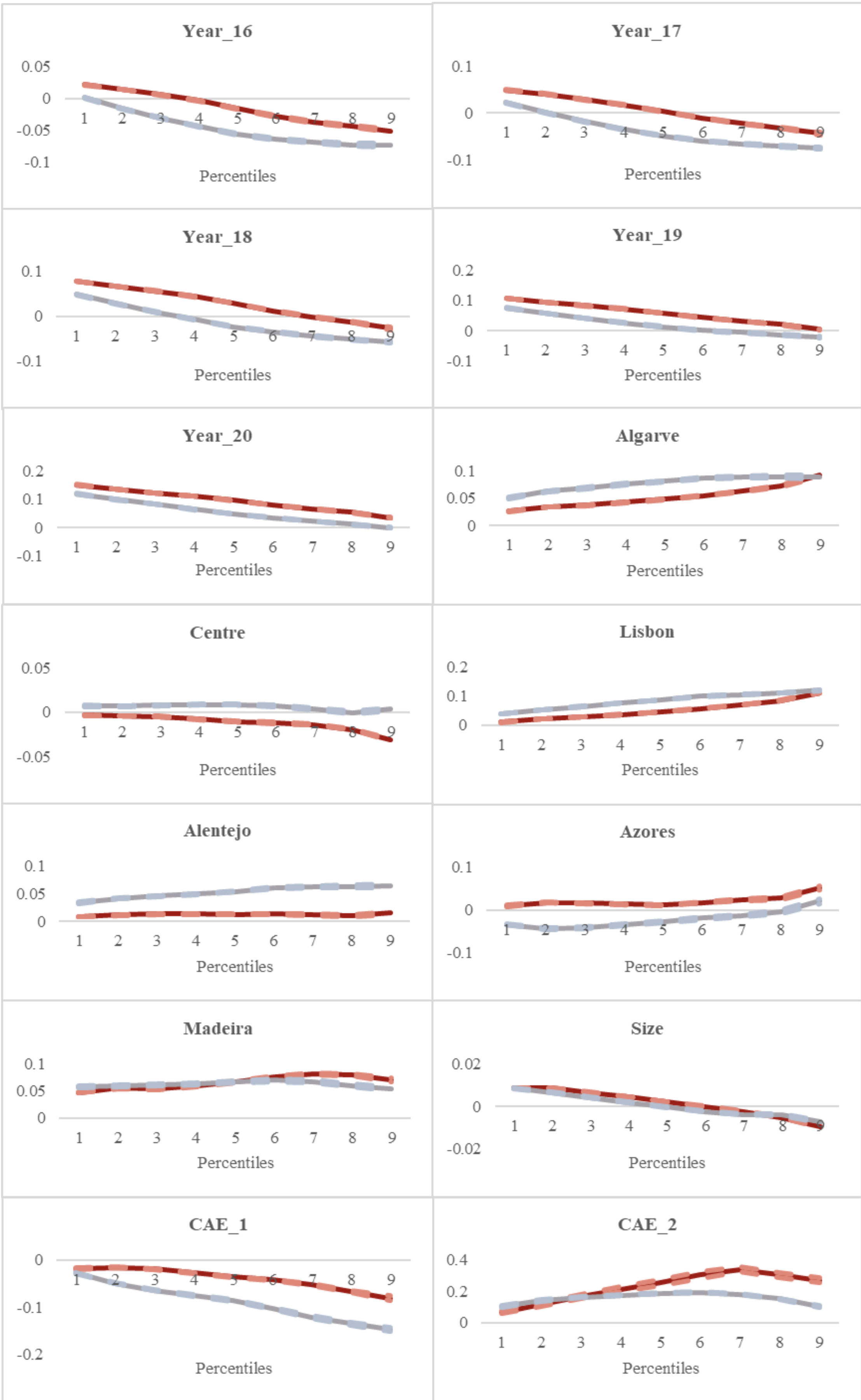
(ii) *, **, *** denote significant at 1%, 5%, and 10%, respectively.

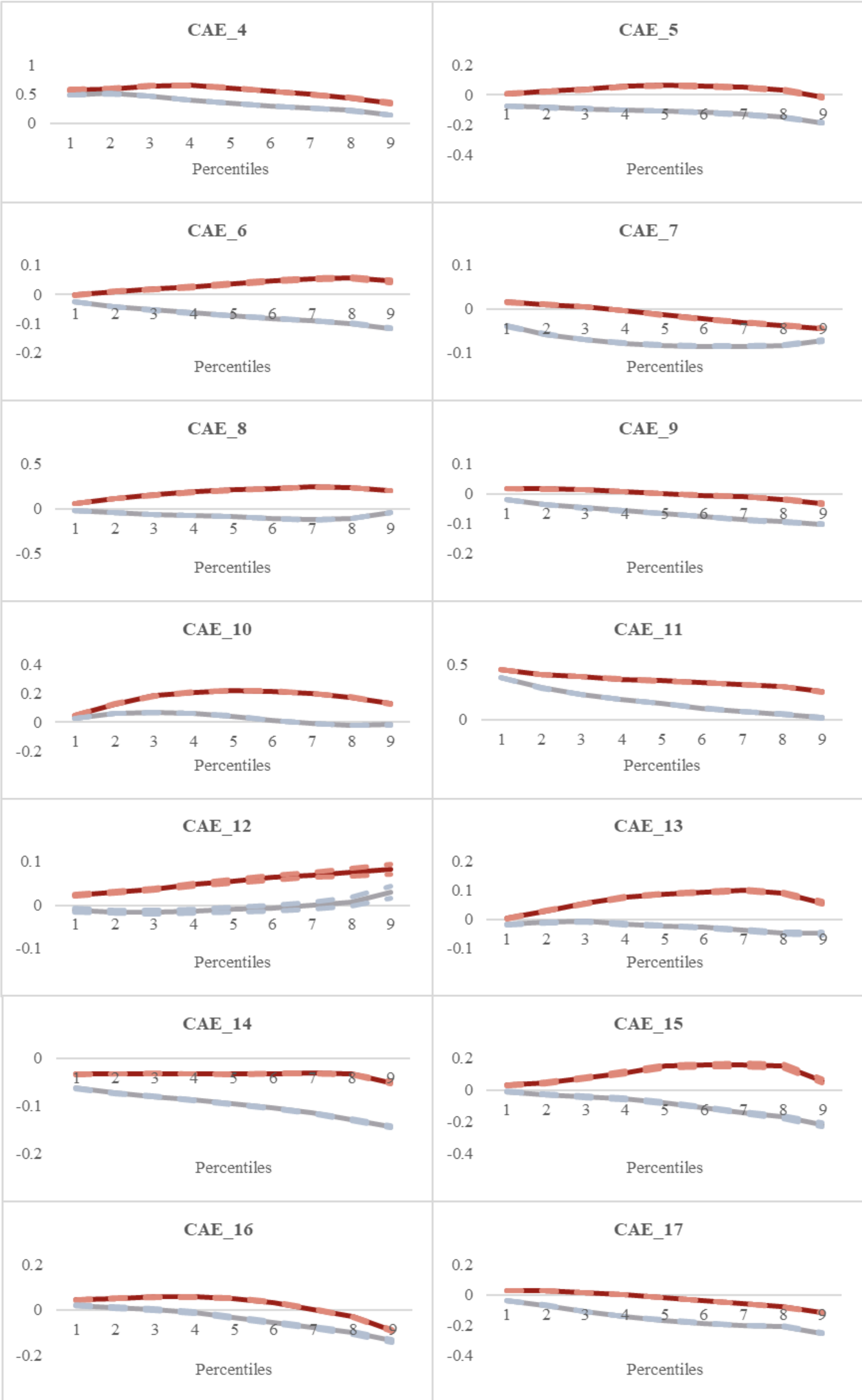
Source: Quadros de Pessoal. Own elaboration.

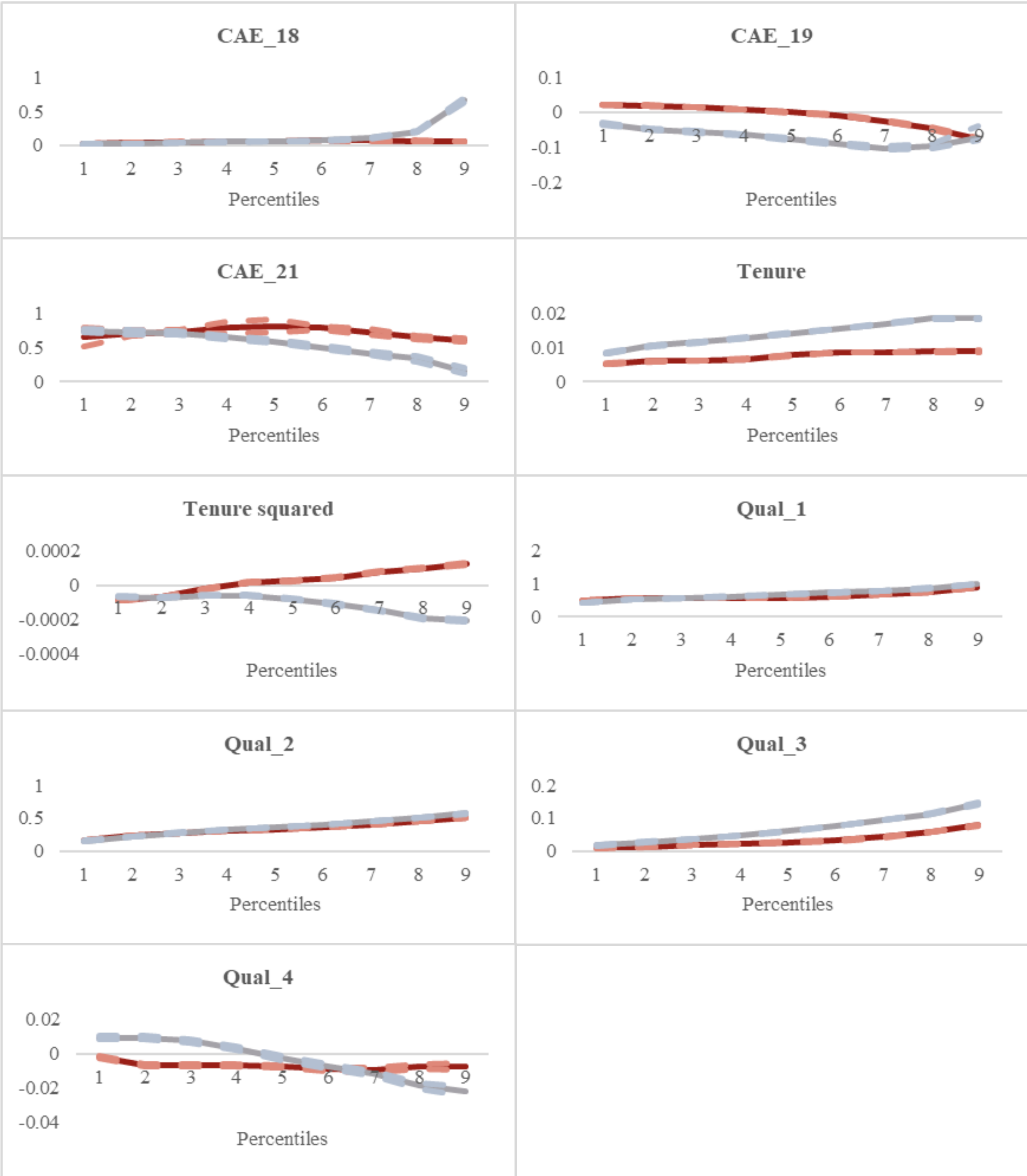
Annex E - Quantile regressions

Figure 10 - Estimated coefficients obtained through Specification *4* quantile regressions




Year_16
0.05
0
-0.05
-0.1
1 2 3 4 5 6 7 8 9
Percentiles
Year_17
0.1
0
-0.1
1 2 3 4 5 6 7 8 9
Percentiles
Year_18
0.1
0
-0.1
1 2 3 4 5 6 7 8 9
Percentiles
Year_19
0.2
0.1
0
-0.1
1 2 3 4 5 6 7 8 9
Percentiles
Year_20
0.2
0.1
0
-0.1
1 2 3 4 5 6 7 8 9
Percentiles
Algarve
0.1
0.05
0
1 2 3 4 5 6 7 8 9
Percentiles
Centre
0.05
0
-0.05
1 2 3 4 5 6 7 8 9
Percentiles
Lisbon
0.2
0.1
0
1 2 3 4 5 6 7 8 9
Percentiles
Alentejo
0.1
0.05
0
1 2 3 4 5 6 7 8 9
Percentiles
Azores
0.1
0
-0.1
1 2 3 4 5 6 7 8 9
Percentiles
Madeira
0.1
0.05
0
1 2 3 4 5 6 7 8 9
Percentiles
Size
0.02
0
-0.02
1 2 3 4 5 6 7 8 9
Percentiles
CAE_1
0
-0.1
-0.2
1 2 3 4 5 6 7 8 9
Percentiles
CAE_2
0.4
0.2
0
1 2 3 4 5 6 7 8 9
Percentiles


CAE_4
1
0.5
0
1 2 3 4 5 6 7 8 9
Percentiles
CAE_5
0.2
0
-0.2
-0.4
1 2 3 4 5 6 7 8 9
Percentiles
CAE_6
0.1
0
-0.1
-0.2
1 2 3 4 5 6 7 8 9
Percentiles
CAE_7
0.1
0
-0.1
1 2 3 4 5 6 7 8 9
Percentiles
CAE_8
0.5
0
-0.5
1 2 3 4 5 6 7 8 9
Percentiles
CAE_9
0.1
0
-0.1
-0.2
1 2 3 4 5 6 7 8 9
Percentiles
CAE_10
0.4
0.2
0
-0.2
1 2 3 4 5 6 7 8 9
Percentiles
CAE_11
0.5
0
1 2 3 4 5 6 7 8 9
Percentiles
CAE_12
0.1
0
-0.1
1 2 3 4 5 6 7 8 9
Percentiles
CAE_13
0.2
0.1
0
-0.1
1 2 3 4 5 6 7 8 9
Percentiles
CAE_14
0
-0.1
-0.2
1 2 3 4 5 6 7 8 9
Percentiles
CAE_15
0.2
0
-0.2
-0.4
1 2 3 4 5 6 7 8 9
Percentiles
CAE_16
0.2
0
-0.2
1 2 3 4 5 6 7 8 9
Percentiles
CAE_17
0.2
0
-0.2
-0.4
1 2 3 4 5 6 7 8 9
Percentiles



Notes: (i) Solid lines represent the estimated coefficient at different deciles.

(ii) Dotted lines represent a 95% confidence interval.

(iii) Red represents the women sample and gray the male.

Source: Quadros de Pessoal. Own elaboration.

Annex F - Gelbach's decomposition

Table 22 - Decomposition of the immigrant-native wage gap variation between specifications 2 and 4

	Women			Men		
	Coefficient	**S.E.**	**%**	**Coefficient**	**S.E.**	**%**
Concentration	-0.0142*	0.000276	11.74	-0.0244*	0.000327	16.83
Size	-0.00184*	0.000119	1.52	-0.00668*	0.000143	4.61
Industry (CAE)	-0.00987*	0.000213	8.16	-0.0165*	0.000219	11.38
Tenure	-0.0476*	0.000217	39.34	-0.0585*	0.000218	40.34
Qualification levels	-0.0471*	0.000501	38.93	-0.0387*	0.000436	26.69
Total	-0.1210*	0.000664	100	-0.1450*	0.000624	100

Notes: (i) S.E. is the standard error.

(ii) % is the percentage of the explained difference and it is calculated by dividing the decomposition coefficient of each covariant by "Total".

(iii) The "Total" line is the sum of all X_{2k} variables' impacts on Immigrant.

(iv) *, **, *** denote significant at 1%, 5%, and 10%, respectively.

Source: Quadros de Pessoal. Own elaboration.

Table 23 - Decomposition of the immigrant-native wage gap variation between specifications 2 and 4 considering time effects

	Women			Men		
	Coefficient	**S.E.**	**%**	**Coefficient**	**S.E.**	**%**
Concentration	-0.0147*	0.000284	12.89	-0.0253*	0.000340	17.57
Size	-0.00182*	0.000118	1.60	-0.00674*	0.000144	4.68
Industry (CAE)	-0.0119*	0.000213	10.44	-0.0181*	0.000219	12.57
Tenure	-0.0493*	0.000220	43.25	-0.0618*	0.000223	42.92
Qualification levels	-0.0490*	0.000500	42.98	-0.0427*	0.000434	29.65
$Year_{12}$	0.00115*	3.30e-05	-1.01	0.00141*	3.57e-05	-0.98
$Year_{13}$	0.00178*	4.01e-05	-1.56	0.00222*	4.27e-05	-1.54
$Year_{14}$	0.00165*	3.91e-05	-1.45	0.00211*	4.22e-05	-1.47
$Year_{15}$	0.00188*	4.31e-05	-1.65	0.00225*	4.52e-05	-1.56
$Year_{16}$	0.00135*	4.00e-05	-1.18	0.00175*	4.51e-05	-1.22
$Year_{17}$	0.000500*	3.24e-05	-0.44	0.000799*	4.40e-05	-0.55
$Year_{18}$	-0.000332*	2.16e-05	0.29	-0.000734*	3.63e-05	0.51
$Year_{19}$	0.000779*	5.63e-05	-0.68	-0.00113*	6.75e-05	0.78
$Year_{20}$	0.00377*	7.99e-05	-3.31	0.00163*	8.76e-05	-1.13
Total	-0.114*	0.000668	100	-0.144*	0.000627	100

Notes: (i) S.E. is the standard error.

(ii) % is the percentage of the explained difference and it is calculated by dividing the decomposition coefficient of each covariant by "Total".

(iii) The "Total" line is the sum of all X_{2k} variables' impacts on Immigrant.

(iv) *, **, *** denote significant at 1%, 5%, and 10%, respectively.

Source: Quadros de Pessoal. Own elaboration.

Table 24 - Decomposition of the immigrant-native wage gap variation between specifications 2 and 4 considering NUTS II effects

	Women			Men		
	Coefficient	**S.E.**	**%**	**Coefficient**	**S.E.**	**%**
Concentration	-0.0155*	0.000300	14.49	-0.0259*	0.000347	18.91
Size	-0.00184*	0.000119	1.72	-0.00675*	0.000145	4.93
Industry (CAE)	-0.00975*	0.000218	9.11	-0.0210*	0.000230	15.33
Tenure	-0.0499***	0.000220	46.64	-0.0610*	0.000221	44.53
Qualification levels	-0.0450*	0.000498	42.06	-0.0384*	0.000432	28.03
Algarve	0.00460*	8.00e-05	-4.30	0.00378*	6.97e-05	-2.76
Centre	0.00138*	3.59e-05	-1.29	0.000240*	2.22e-05	-0.18
Lisbon	0.00976*	0.000125	-9.12	0.00927*	0.000129	-6.77
Alentejo	0.000153*	1.23e-05	-0.14	0.00279*	5.59e-05	-2.04
Azores	-0.000256*	1.43e-05	0.24	7.71e-05*	1.47e-05	-0.06
Madeira	-0.000679*	2.28e-05	0.63	-0.000545*	1.84e-05	0.40
Total	-0.107*	0.000678	100	-0.137*	0.000629	100

Notes: (i) S.E. is the standard error.

(ii) % is the percentage of the explained difference and it is calculated by dividing the decomposition coefficient of each covariant by "Total".

(iii) The "Total" line is the sum of all X_{2k} variables' impacts on Immigrant.

(iv) *, **, *** denote significant at 1%, 5%, and 10%, respectively.

Source: Quadros de Pessoal. Own elaboration.

Table 25 - Decomposition of the immigrant-native wage gap variation between specifications 2 and 4 considering group of nationality effects

	Women			**Men**		
	Coefficient	**S.E.**	**%**	**Coefficient**	**S.E.**	**%**
Concentration	-0.0159*	0.000278	6.60	-0.0262*	0.000329	7.87
Size	-0.00104*	0.000119	0.43	-0.00529*	0.000143	1.59
Industry (CAE)	-0.00985*	0.000214	4.09	-0.0158*	0.000220	4.74
Tenure	-0.0474*	0.000216	19.67	-0.0583*	0.000218	17.51
Qualification levels	-0.0469*	0.000500	19.46	-0.0383*	0.000433	11.50
Brazil	-0.0591*	0.000808	24.52	-0.0630*	0.000891	18.92
CEEC	-0.0211*	0.000335	8.76	-0.0195*	0.000308	5.86
China	-0.00370*	0.000115	1.54	-0.00819*	0.000150	2.46
PALOP	-0.0191*	0.000331	7.93	-0.0304*	0.000409	9.13
SA	-0.00432*	0.000143	1.79	-0.0465*	0.000603	13.96
Others	-0.0122*	0.000293	5.06	-0.0216*	0.000433	6.49
Total	-0.241*	0.00178	100	-0.333*	0.00238	100

Notes: (i) S.E. is the standard error.

(ii) % is the percentage of the explained difference and it is calculated by dividing the decomposition coefficient of each covariant by "Total".

(iii) The "Total" line is the sum of all X_{2k} variables' impacts on Immigrant.

(iv) *, **, *** denote significant at 1%, 5%, and 10%, respectively.

Source: Quadros de Pessoal. Own elaboration.

Annex G - Variable selection results

Table 26 - Estimated coefficients resulting from applying high-dimensionality reduction procedure – gender samples (2011 - 2020)

(Dependent variable: log of the real hourly wage)

	Full sample	Female sample	Male Sample
Immigrant	0.503*	0.172*	0.565*
	(0.00597)	(0.00674)	(0.00853)
YSM	-0.00193*	0.000336*	-0.00196*
	(0.000179)	(0.000111)	(0.000243)
YSM^2	0.000116*	-	0.000124*
	(7.32e-06)		(9.58e-06)
Age	0.0204*	0.00553*	0.0249*
	(7.83e-05)	(1.78e-05)	(0.000113)
Age^2	-0.000172*	-	-0.000217*
	(9.69e-07)		(1.39e-06)
$Educ_2$	0.145*	0.155*	0.145*
	(0.000267)	(0.000337)	(0.000386)
$Educ_3$	0.380*	0.399*	0.411*
	(0.000470)	(0.000588)	(0.000735)
$Educ_4$	0.399*	0.431*	0.411*
	(0.000976)	(0.00129)	(0.00143)
$Educ_5$	0.711*	0.749*	0.694*
	(0.00266)	(0.00330)	(0.00413)
$Year_{12}$	-0.0355*	-0.0348*	-0.0352*
	(0.000521)	(0.000671)	(0.000739)
$Year_{13}$	-0.0480*	-0.0445*	-0.0497*
	(0.000525)	(0.000677)	(0.000747)
$Year_{14}$	-0.0483*	-0.0433*	-0.0513*
	(0.000522)	(0.000673)	(0.000743)
$Year_{15}$	-0.0555*	-0.0495*	-0.0595*
	(0.000517)	(0.000666)	(0.000738)
$Year_{16}$	-0.0555*	-0.0468*	-0.0624*
	(0.000511)	(0.000660)	(0.000730)
$Year_{17}$	-0.0516*	-0.0404*	-0.0617*
	(0.000502)	(0.000648)	(0.000718)

$Year_{18}$	-0.0353*	-0.0251*	-0.0448*
	(0.000495)	(0.000638)	(0.000709)
$Year_{19}$	-0.00918*	-0.000840	-0.0174*
	(0.000494)	(0.000634)	(0.000708)
$Year_{20}$	0.0139*	0.0236*	0.00368*
	(0.000492)	(0.000636)	(0.000703)
Algarve	0.0666*	0.0573*	0.0674*
	(0.000623)	(0.000751)	(0.000966)
Centre	-0.0213*	-0.0341*	-0.00554*
	(0.000352)	(0.000430)	(0.000507)
Lisbon	0.0934*	0.0765*	0.101*
	(0.000298)	(0.000381)	(0.000434)
Alentejo	0.0378*	0.00866*	0.0603*
	(0.000594)	(0.000697)	(0.000869)
Azores	0.00537*	0.0294*	-0.00702*
	(0.00113)	(0.00141)	(0.00167)
Madeira	0.0723*	0.0754*	0.0670*
	(0.000840)	(0.00104)	(0.00126)
Concentration	-0.0114*	-0.00833*	-0.0112*
	(0.000123)	(0.000163)	(0.000175)
Size	-0.000329*	0.000413*	0.00402*
	(7.75e-05)	(9.70e-05)	(0.000118)
CAE_1	-0.109*	-0.0883*	-0.110*
	(0.00112)	(0.00142)	(0.00153)
CAE_2	0.205*	0.195*	0.132*
	(0.00182)	(0.00658)	(0.00186)
CAE_4	0.371*	0.434*	0.292*
	(0.00122)	(0.00327)	(0.00133)
CAE_5	-0.0706*	0.0165*	-0.139*
	(0.000903)	(0.00204)	(0.00102)
CAE_6	-0.0285*	-0.00880*	-0.0886*
	(0.000565)	(0.00173)	(0.000623)
CAE_7	-0.0742*	-0.0271*	-0.0696*
	(0.000391)	(0.000510)	(0.000561)
CAE_8	0.00275*	0.135*	-0.0846*
	(0.000538)	(0.00108)	(0.000630)
CAE_9	-0.0939*	-0.0347*	-0.0891*
	(0.000480)	(0.000579)	(0.000764)

CAE_{10}	0.0366* (0.000739)	0.109* (0.00115)	-0.0192* (0.000954)
CAE_{11}	0.128* (0.000569)	0.232* (0.000775)	0.0583* (0.000812)
CAE_{12}	-0.00391 (0.00246)	0.0457* (0.00302)	-
CAE_{13}	-0.0226* (0.000864)	0.0388* (0.00119)	-0.0480* (0.00121)
CAE_{14}	-0.114* (0.000473)	-0.0804* (0.000669)	-0.137* (0.000643)
CAE_{15}	-0.0162* (0.00226)	0.110* (0.00363)	-0.0880* (0.00287)
CAE_{16}	-0.121* (0.000942)	-0.0236* (0.00104)	-0.113* (0.00201)
CAE_{17}	-0.178* (0.000378)	-0.0676* (0.000451)	-0.185* (0.000895)
CAE_{18}	0.0832* (0.00211)	0.0250* (0.00198)	0.129* (0.00315)
CAE_{19}	-0.116* (0.000972)	-0.0301* (0.00104)	-0.0868* (0.00213)
CAE_{21}	0.484* (0.0162)	0.580* (0.0256)	0.485* (0.0184)
Tenure	0.0124* (4.06e-05)	0.00874* (2.28e-05)	0.0154* (5.75e-05)
$Tenure^2$	-0.000105* (1.31e-06)	0.628* (0.000916)	-0.000180* (1.80e-06)
$Qual_1$	0.678* (0.000712)	0.353* (0.000796)	0.675* (0.00105)
$Qual_2$	0.369* (0.000613)	0.0471* (0.000624)	0.362* (0.000886)
$Qual_3$	0.0556* (0.000495)	-0.00761* (0.000669)	0.0626* (0.000733)
$Qual_4$	-0.0146* (0.000536)	-0.0347* (0.000579)	-0.0279* (0.000803)
Brazil	-0.217* (0.00203)	-0.197* (0.00227)	-0.232* (0.00307)
CEEC	-0.228* (0.00202)	-0.200* (0.00235)	-0.246* (0.00305)

China	-0.231* (0.00269)	-0.167* (0.00341)	-0.278* (0.00393)
PALOP	-0.243* (0.00200)	-0.201* (0.00234)	-0.272* (0.00301)
SA	-0.210* (0.00201)	-0.169* (0.00294)	-0.248* (0.00295)
Others	-0.180* (0.00247)	-0.153* (0.00285)	-0.210* (0.00360)
Age · Immig	-0.0149* (0.000283)	0.000941* (0.000322)	-0.0167* (0.000398)
Age^2 · Immig	0.000122* (3.63e-06)	-6.85e-05* (4.19e-06)	0.000135* (5.08e-06)
$Educ_2$ · Immig	-0.108* (0.000867)	-0.107* (0.000969)	-0.107* (0.00127)
$Educ_3$ · Immig	-0.140* (0.00213)	-0.147* (0.00234)	-0.133* (0.00341)
$Educ_4$ · Immig	-0.121* (0.00545)	-0.104* (0.00635)	-0.113* (0.00874)
$Educ_5$ · Immig	-0.128* (0.00890)	0.0113* (0.00233)	-0.0923* (0.0122)
$Year_{12}$ · Immig	0.00853* (0.00201)	0.0177* (0.00241)	0.00619** (0.00288)
$Year_{13}$ · Immig	0.0128* (0.00206)	0.0331* (0.00240)	0.00987* (0.00296)
$Year_{14}$ · Immig	0.0268* (0.00202)	0.0368* (0.00240)	0.0222* (0.00290)
$Year_{15}$ · Immig	0.0289* (0.00203)	0.0516* (0.00234)	0.0223* (0.00291)
$Year_{16}$ · Immig	0.0432* (0.00196)	0.0685* (0.00225)	0.0388* (0.00279)
$Year_{17}$ · Immig	0.0595* (0.00189)	-0.197* (0.00227)	0.0563* (0.00269)
$Year_{18}$ · Immig	0.0679* (0.00180)	0.0788* (0.00214)	0.0647* (0.00256)
$Year_{19}$ · Immig	0.0683* (0.00172)	0.0842* (0.00205)	0.0609* (0.00245)

$Year_{20}$ · Immig	0.0925*	0.103*	0.0902*
	(0.00170)	(0.00206)	(0.00241)
Algarve · Immig	-0.0439*	0.00499**	-0.0604*
	(0.00186)	(0.00214)	(0.00270)
Centre · Immig	1.75e-05	0.0275*	-0.0192*
	(0.00157)	(0.00190)	(0.00217)
Lisbon · Immig	-0.0637*	-0.0354*	-0.0724*
	(0.00156)	(0.00177)	(0.00219)
Alentejo · Immig	-0.0468*	-0.00609**	-0.0687*
	(0.00192)	(0.00251)	(0.00261)
Azores · Immig	-0.0231*	-0.00656	-0.0376*
	(0.00774)	(0.00897)	(0.0108)
Madeira · Immig	-0.00105	-0.000371	0.00625
	(0.00588)	(0.00577)	(0.00886)
Concentration · Immig	-0.0148*	-0.0105*	-0.0209*
	(0.000461)	(0.000598)	(0.000639)
Size · Immig	0.00819*	0.00670*	0.00456*
	(0.000268)	(0.000332)	(0.000389)
CAE_1 · Immig	0.0641*	0.0874*	0.0570*
	(0.00191)	(0.00269)	(0.00251)
CAE_2 · Immig	-0.0330*	-	-0.00712
	(0.0109)		(0.0112)
CAE_4 · Immig	-0.0910*	-	-0.149*
	(0.0214)		(0.0223)
CAE_5 · Immig	0.0117*	-0.0456*	0.0502*
	(0.00305)	(0.00571)	(0.00354)
CAE_6 · Immig	0.00954*	0.0114**	0.0384*
	(0.00154)	(0.00561)	(0.00182)
CAE_7 · Immig	0.0388*	0.0323*	0.0355*
	(0.00156)	(0.00206)	(0.00214)
CAE_8 · Immig	0.00323	-0.0124***	0.0431*
	(0.00229)	(0.00680)	(0.00253)
CAE_9 · Immig	0.0532*	0.0544*	0.0424*
	(0.00137)	(0.00178)	(0.00200)
CAE_{10} · Immig	-0.0361*	-0.00512	-0.0440*
	(0.00410)	(0.00603)	(0.00545)
CAE_{11} · Immig	0.105*	0.0493*	0.183*
	(0.00587)	(0.00642)	(0.0100)

CAE_{12} · Immig	-0.0223*	-0.0146**	-0.0257*
	(0.00509)	(0.00609)	(0.00732)
CAE_{13} · Immig	0.00728***	-0.0161*	0.0285*
	(0.00374)	(0.00461)	(0.00547)
CAE_{14} · Immig	-0.000770	0.0315*	0.00170
	(0.00163)	(0.00217)	(0.00221)
CAE_{15} · Immig	-0.0640*	-0.0645*	-0.0605*
	(0.0128)	(0.0199)	(0.0175)
CAE_{16} · Immig	0.0873*	0.117*	0.00439
	(0.00427)	(0.00499)	(0.00769)
CAE_{17} · Immig	0.0622*	0.0499*	0.0486*
	(0.00179)	(0.00203)	(0.00510)
CAE_{18} · Immig	0.560*	0.0824*	0.763*
	(0.0134)	(0.00722)	(0.0182)
CAE_{19} · Immig	0.0433*	0.0313*	0.0106***
	(0.00277)	(0.00296)	(0.00598)
CAE_{21} · Immig	0.294*	0.287*	-
	(0.0408)	(0.0508)	
Tenure · Immig	-0.00175*	-	-0.00244*
	(0.000125)		(0.000173)
$Tenure^2$ · Immig	-	-	-
$Qual_1$ · Immig	0.0960*	0.000319	0.144*
	(0.00346)	(0.00426)	(0.00495)
$Qual_2$ · Immig	0.0459*	-0.0380*	0.0872*
	(0.00281)	(0.00309)	(0.00391)
$Qual_3$ · Immig	0.00651*	-0.00570*	0.00214
	(0.00109)	(0.00141)	(0.00156)
$Qual_4$ · Immig	0.0237*	0.00816*	0.0410*
	(0.00117)	(0.00153)	(0.00168)
Constant	0.656*	0.869*	0.586*
	(0.00161)	(0.00114)	(0.00231)
N	8,870,100	4,078,776	4,791,324

Notes: (i) In parentheses are the robust standard errors.

(ii) *, **, *** denote significant at 1%, 5%, and 10%, respectively.

(iii) Immig is the "Immigrant" variable.

(iv) The above results define the optimal number of control variables assuming the exogeneity of the variable of interest.

Source: Quadros de Pessoal. Own elaboration.

Table 27 - Estimated coefficients resulting from applying high-dimensionality reduction procedure – NUTS II samples (2011 - 2020)

(Dependent variable: log of the real hourly wage)

	North	Algarve	Centre	Lisbon	Alentejo	Azores	Madeira
Immigrant	0.282*	0.149*	0.122*	0.529*	0.363*	0.283*	0.466*
	(0.0193)	(0.0142)	(0.0116)	(0.00842)	(0.0175)	(0.0947)	(0.0670)
YSM	-0.000680*	0.00264*	0.00204*	-0.00228*	0.00133*	-	-0.00429*
	(0.000257)	(0.000226)	(0.000193)	(0.000247)	(0.000339)		(0.00111)
YSM^2	-	-	-	0.000114*	-	-	-
				(9.97e-06)			
Age	0.00665*	-	0.00406*	0.0236*	0.00416*	0.00708*	0.00588*
	(2.70e-05)		(2.96e-05)	(0.000119)	(5.98e-05)	(0.000149)	(0.000105)
Age^2	-	-	-	-0.000203*	-	-	-
				(1.47e-06)			
$Educ_2$	0.154*	0.0887*	0.0995*	0.150*	0.0764*	0.169*	0.122*
	(0.000525)	(0.00130)	(0.000652)	(0.000386)	(0.00122)	(0.00260)	(0.00208)
$Educ_3$	0.368*	0.248*	0.361*	0.393*	0.335*	0.508*	0.292*
	(0.000977)	(0.00256)	(0.00133)	(0.000627)	(0.00273)	(0.00579)	(0.00375)
$Educ_4$	0.385*	0.298*	0.351*	0.417*	0.384*	0.696*	0.337*
	(0.00196)	(0.00752)	(0.00284)	(0.00127)	(0.00675)	(0.0188)	(0.0119)
$Educ_5$	0.766*	0.682*	0.802*	0.628*	0.420*	0.847*	0.354*
	(0.00416)	(0.0293)	(0.00716)	(0.00368)	(0.0350)	(0.0513)	(0.0472)
$Year_{12}$	-0.0295*	-0.0156*	-0.0331*	-0.0392*	-0.0146*	-	-
	(0.00100)	(0.00216)	(0.00127)	(0.000750)	(0.00202)		
$Year_{13}$	-0.0360*	-0.0326*	-0.0416*	-0.0561*	-0.0142*	-0.0457*	-
	(0.00101)	(0.00222)	(0.00126)	(0.000755)	(0.00208)	(0.00362)	
$Year_{14}$	-0.0322*	-0.0281*	-0.0358*	-0.0591*	-0.0160*	-0.0444*	-0.0224*
	(0.00101)	(0.00212)	(0.00125)	(0.000752)	(0.00199)	(0.00360)	(0.00286)
$Year_{15}$	-0.0371*	-0.0315*	-0.0346*	-0.0693*	-	-0.0690*	-0.0191*
	(0.000996)	(0.00203)	(0.00125)	(0.000747)		(0.00341)	(0.00286)
$Year_{16}$	-0.0338*	-0.0223*	-0.0268*	-0.0741*	-0.0154*	-0.0665*	-
	(0.000975)	(0.00196)	(0.00123)	(0.000743)	(0.00191)	(0.00354)	
$Year_{17}$	-0.0320*	-	-0.0187*	-0.0715*	-	-0.0397*	-
	(0.000956)		(0.00120)	(0.000734)		(0.00362)	
$Year_{18}$	-0.0162*	0.00829*	0.00781*	-0.0600*	0.0113*	-	-
	(0.000935)	(0.00185)	(0.00118)	(0.000729)	(0.00175)		
$Year_{19}$	0.00805*	0.0360*	0.0327*	-0.0318*	0.0325*	0.0124*	-0.0186*
	(0.000931)	(0.00183)	(0.00117)	(0.000729)	(0.00182)	(0.00367)	(0.00222)

Year$_{20}$	0.0323*	0.0675*	0.0577*	-0.0153*	0.0729*	0.0643*	-
	(0.000931)	(0.00189)	(0.00115)	(0.000731)	(0.00174)	(0.00376)	
Concentration	0.0153*	0.00150**	0.00635*	-0.0177*	-0.0408*	0.00779*	0.00295**
	(0.000278)	(0.000636)	(0.000305)	(0.000178)	(0.000625)	(0.00179)	(0.00130)
Size	0.0136*	0.0372*	0.0166*	-0.00744*	0.0122*	0.0102*	0.0185*
	(0.000157)	(0.000419)	(0.000266)	(0.000104)	(0.000496)	(0.00169)	(0.000862)
CAE$_{1}$	-0.111*	-0.0376*	-0.118*	-0.188*	-0.0211*	0.113*	0.0990*
	(0.00298)	(0.00415)	(0.00191)	(0.00264)	(0.00242)	(0.00664)	(0.0207)
CAE$_{2}$	-0.0483*	0.128*	0.119*	-0.0473*	0.303*	-	0.392*
	(0.00623)	(0.0179)	(0.00314)	(0.00440)	(0.00236)		(0.0294)
CAE$_{4}$	0.182*	-	0.450*	0.307*	-	0.388*	0.757*
	(0.0205)		(0.00809)	(0.00139)		(0.00421)	(0.0222)
CAE$_{5}$	-0.0228*	0.0473*	-0.0234*	-0.134*	0.0392*	0.0544*	-0.0843*
	(0.00218)	(0.00335)	(0.00274)	(0.00127)	(0.00504)	(0.00900)	(0.00613)
CAE$_{6}$	-0.0191*	0.0534*	-0.0152*	-0.0708*	-0.0142*	-0.0365*	0.0355*
	(0.00100)	(0.00319)	(0.00112)	(0.00102)	(0.00268)	(0.00361)	(0.00431)
CAE$_{7}$	-0.0822*	-0.0170*	-0.0364*	-0.108*	0.0475*	-0.0881*	-0.0868*
	(0.000637)	(0.00278)	(0.000999)	(0.000703)	(0.00218)	(0.00330)	(0.00423)
CAE$_{8}$	-0.0346*	-0.100*	-0.0660*	-0.0220*	-0.0854*	0.379*	0.200*
	(0.00118)	(0.00332)	(0.00104)	(0.000842)	(0.00227)	(0.00716)	(0.00539)
CAE$_{9}$	-0.0869*	0.0937*	-0.0735*	-0.172*	-0.0321*	-0.0474*	-0.0196*
	(0.000931)	(0.00263)	(0.00129)	(0.000787)	(0.00257)	(0.00341)	(0.00375)
CAE$_{10}$	0.0824*	0.104*	0.0307*	-0.0174*	-0.0202**	-	0.148*
	(0.00178)	(0.0127)	(0.00308)	(0.000940)	(0.0102)		(0.0117)
CAE$_{11}$	0.158*	0.465*	0.316*	0.0615*	0.326*	0.328*	0.126*
	(0.00111)	(0.00787)	(0.00459)	(0.000824)	(0.00746)	(0.00560)	(0.00534)
CAE$_{12}$	-0.0108*	0.114*	-0.126*	-0.0284*	-	-	-0.0736*
	(0.00189)	(0.00455)	(0.00801)	(0.00372)			(0.00874)
CAE$_{13}$	-0.0550*	0.0633*	-0.0773*	-0.0644*	-	-	0.193*
	(0.00120)	(0.00655)	(0.00307)	(0.00111)			(0.00798)
CAE$_{14}$	-0.0162*	-0.0208*	-0.0823*	-0.150*	-0.0508*	-0.0465*	-0.105*
	(0.000935)	(0.00347)	(0.00179)	(0.000730)	(0.00272)	(0.00900)	(0.00539)
CAE$_{15}$	0.0435*	-0.0148**	-0.111*	-0.0494*	-0.0980*	0.153*	-
	(0.00748)	(0.00586)	(0.00373)	(0.00299)	(0.00603)	(0.0145)	
CAE$_{16}$	-0.0762*	0.0703*	-0.0497*	-0.193*	-	-0.0649*	-0.211*
	(0.00194)	(0.00570)	(0.00279)	(0.00126)		(0.00817)	(0.00893)
CAE$_{17}$	-0.151*	-0.125*	-0.137*	-0.236*	-0.186*	-0.00282	-0.0870*
	(0.000694)	(0.00274)	(0.000718)	(0.000714)	(0.00141)	(0.00375)	(0.00435)

CAE_{18}	0.136*	0.147*	-	0.0208*	0.0834*	0.100*	0.146*
	(0.00432)	(0.00461)		(0.00303)	(0.0167)	(0.0127)	(0.0105)
CAE_{19}	-0.116*	-0.0347*	-0.0813*	-0.163*	-0.105*	-	-0.109*
	(0.00198)	(0.00458)	(0.00197)	(0.00139)	(0.00425)		(0.00911)
CAE_{21}	-	0.00769*	-	0.419*	-	-	-
		(8.00e-05)		(0.0161)			
Tenure	0.0112*	-0.0148**	0.00732*	0.0149*	0.00985*	0.00719*	0.00404*
	(7.58e-05)	(0.00586)	(4.00e-05)	(5.93e-05)	(8.70e-05)	(0.000163)	(0.000123)
$Tenure^2$	-0.000163*	-	-	-0.000134*	-	-	-
	(2.43e-06)			(1.83e-06)			
$Qual_1$	0.702*	0.659*	0.625*	0.710*	0.701*	0.444*	0.757*
	(0.00130)	(0.00387)	(0.00183)	(0.00114)	(0.00362)	(0.00744)	(0.00451)
$Qual_2$	0.381*	0.394*	0.369*	0.399*	0.429*	0.190*	0.405*
	(0.00110)	(0.00315)	(0.00141)	(0.00103)	(0.00266)	(0.00583)	(0.00358)
$Qual_3$	0.0510*	0.104*	0.0818*	0.0868*	0.120*	0.0149*	0.137*
	(0.000770)	(0.00235)	(0.00100)	(0.000923)	(0.00183)	(0.00464)	(0.00250)
$Qual_4$	-0.0355*	0.00589**	-0.0200*	0.0310*	0.0283*	-0.0562*	0.00102
	(0.000848)	(0.00251)	(0.00108)	(0.000987)	(0.00199)	(0.00485)	(0.00298)
Brazil	-0.180*	-0.228*	-0.158*	-0.261*	-0.160*	-0.181*	-0.206*
	(0.00513)	(0.00520)	(0.00588)	(0.00283)	(0.00939)	(0.0443)	(0.0262)
CEEC	-0.221*	-0.232*	-0.179*	-0.268*	-0.176*	-0.254*	-0.320*
	(0.00487)	(0.00496)	(0.00573)	(0.00296)	(0.00918)	(0.0432)	(0.0206)
China	-0.176*	-0.275*	-0.189*	-0.272*	-0.161*	-0.207*	-0.333*
	(0.00634)	(0.00657)	(0.00682)	(0.00411)	(0.0105)	(0.0420)	(0.0239)
PALOP	-0.213*	-0.258*	-0.189*	-0.277*	-0.163*	-0.224*	-0.286*
	(0.00518)	(0.00535)	(0.00599)	(0.00281)	(0.0100)	(0.0438)	(0.0252)
SA	-0.169*	-0.238*	-0.190*	-0.243*	-0.168*	-0.124**	-0.252*
	(0.00669)	(0.00539)	(0.00610)	(0.00289)	(0.00916)	(0.0528)	(0.0271)
Others	-0.163*	-0.221*	-0.175*	-0.170*	-0.171*	-0.222*	-0.266*
	(0.00651)	(0.00607)	(0.00636)	(0.00371)	(0.00984)	(0.0415)	(0.0217)
Age · Immig	-0.00936*	0.00590*	0.00262*	-0.0174*	0.000430	-0.00589	-0.00892*
	(0.000943)	(0.000633)	(0.000494)	(0.000408)	(0.000719)	(0.00451)	(0.00342)
Age^2 · Immig	6.89e-05*	-0.000109*	-7.82e-05*	0.000151*	-5.20e-05*	-3.03e-07	4.05e-05
	(1.19e-05)	(8.06e-06)	(6.32e-06)	(5.29e-06)	(9.56e-06)	(5.72e-05)	(4.37e-05)

$Educ_2$ · Immig	-0.115*	-0.0593*	-0.0768*	-0.0952*	-0.0545*	-0.173*	-0.123*
	(0.00303)	(0.00252)	(0.00174)	(0.00121)	(0.00272)	(0.0159)	(0.0117)
$Educ_3$ · Immig	-0.139*	-0.0611*	-0.154*	-0.134*	-0.0731*	-0.316*	-0.137*
	(0.00595)	(0.00748)	(0.00551)	(0.00273)	(0.0115)	(0.0390)	(0.0201)
$Educ_4$ · Immig	-0.131*	-	-0.0577*	-0.124*	-0.170*	-	-
	(0.0137)		(0.0179)	(0.00648)	(0.0357)		
$Educ_5$ · Immig	-0.162*	0.240*	-0.0152	0.0158*	-	-	0.558*
	(0.0168)	(0.0497)	(0.0222)	(0.00276)			(0.107)
$Year_{12}$ · Immig	0.00500	-0.0199*	-0.00137	0.0232*	-0.0211*	-0.0170	-0.0541**
	(0.00675)	(0.00535)	(0.00405)	(0.00282)	(0.00648)	(0.0274)	(0.0245)
$Year_{13}$ · Immig	0.00557	-0.0169*	0.00501	0.0382*	-0.0298*	0.00510	-0.0530**
	(0.00691)	(0.00546)	(0.00421)	(0.00278)	(0.00649)	(0.0270)	(0.0267)
$Year_{14}$ · Immig	0.01000	-0.00194	0.0143*	0.0403*	-0.0195*	0.0543**	-0.0191
	(0.00676)	(0.00543)	(0.00411)	(0.00277)	(0.00635)	(0.0267)	(0.0275)
$Year_{15}$ · Immig	0.0153**	-0.00326	0.0180*	0.0576*	-0.0350*	0.0153	-0.0521***
	(0.00693)	(0.00538)	(0.00430)	(0.00268)	(0.00597)	(0.0283)	(0.0267)
$Year_{16}$ · Immig	0.0237*	0.00514	0.0249*	0.0775*	0.00464	0.0462	-0.0621**
	(0.00685)	(0.00511)	(0.00408)	(0.00259)	(0.00595)	(0.0285)	(0.0272)
$Year_{17}$ · Immig	0.0398*	0.00525	0.0308*	0.0910*	0.00960***	0.0464	-0.0621**
	(0.00660)	(0.00451)	(0.00389)	(0.00248)	(0.00546)	(0.0283)	(0.0255)
$Year_{18}$ · Immig	0.0232*	0.0217*	0.0367*	0.0896*	0.0398*	0.0780**	-0.0682*
	(0.00604)	(0.00461)	(0.00370)	(0.00236)	(0.00548)	(0.0311)	(0.0244)
$Year_{19}$ · Immig	0.0242*	0.0247*	0.0382*	0.122*	0.0479*	0.0881*	-0.0365
	(0.00570)	(0.00450)	(0.00353)	(0.00235)	(0.00547)	(0.0316)	(0.0231)
$Year_{20}$ · Immig	0.0392*	0.0530*	0.0602*	-0.0179*	0.0625*	0.0281	-0.0368
	(0.00561)	(0.00467)	(0.00344)	(0.000649)	(0.00532)	(0.0264)	(0.0240)
Concentration · Immig	-0.00391*	-0.0175*	-0.0235*	0.00551*	0.00135	-	-
	(0.00151)	(0.00162)	(0.000916)	(0.000361)	(0.00165)		
Size · Immig	0.0229*	0.00346*	-0.00233*	0.128*	-0.0116*	0.0389*	0.0231*
	(0.00100)	(0.000919)	(0.000699)	(0.00422)	(0.00101)	(0.00526)	(0.00326)
CAE_1 · Immig	0.0398*	0.0137**	0.0780*	-0.0174*	0.00765	-	-0.129*
	(0.00682)	(0.00632)	(0.00315)	(0.000408)	(0.00490)		(0.0498)
CAE_2 · Immig	0.0908*	-0.0640**	-0.0160	0.0642*	0.191*	-	-
	(0.0173)	(0.0294)	(0.0123)	(0.0153)	(0.0335)		

CAE$_4$ · Immig	-	-0.0533	-	-0.0896*	-0.214*	-0.228*	-
		(0.0550)		(0.0236)	(0.0690)	(0.0555)	
CAE$_5$ · Immig	-0.0498*	-0.0641*	0.0511*	0.0404*	-0.0668*	0.00286	-
	(0.0104)	(0.00744)	(0.00910)	(0.00416)	(0.0121)	(0.0550)	
CAE$_6$ · Immig	-0.0313*	-0.0270*	0.0172*	0.0403*	-0.0138**	-0.0518**	0.0208
	(0.00389)	(0.00557)	(0.00276)	(0.00253)	(0.00540)	(0.0209)	(0.0162)
CAE$_7$ · Immig	0.0176*	0.00831	0.00189	0.0703*	-0.0860*	0.111*	0.137*
	(0.00411)	(0.00543)	(0.00296)	(0.00245)	(0.00562)	(0.0200)	(0.0193)
CAE$_8$ · Immig	-0.0167*	0.0599*	0.0440*	0.0516*	0.0473*	-0.326*	0.247*
	(0.00525)	(0.00965)	(0.00322)	(0.00387)	(0.00707)	(0.0589)	(0.0447)
CAE$_9$ · Immig	0.0305*	-0.0442*	0.0432*	0.106*	-0.0199*	0.0251	0.0454*
	(0.00367)	(0.00495)	(0.00285)	(0.00216)	(0.00568)	(0.0181)	(0.0133)
CAE$_{10}$ · Immig	-0.0758*	0.00776	0.0814*	-0.0319*	0.158*	0.320*	0.0981**
	(0.00892)	(0.0390)	(0.0174)	(0.00507)	(0.0429)	(0.0940)	(0.0480)
CAE$_{11}$ · Immig	-0.0224	-0.253*	-	0.157*	-0.0327	-	0.388*
	(0.0142)	(0.0460)		(0.00689)	(0.0573)		(0.0702)
CAE$_{12}$ · Immig	-0.129*	-0.101*	0.0514*	0.0249*	-0.00290	0.120**	0.0907**
	(0.0227)	(0.00891)	(0.0141)	(0.00745)	(0.0176)	(0.0574)	(0.0412)
CAE$_{13}$ · Immig	-0.0449*	-0.119*	0.0190	0.0369*	-0.0444	0.118**	-0.0168
	(0.0100)	(0.0156)	(0.0122)	(0.00456)	(0.0355)	(0.0601)	(0.0395)
CAE$_{14}$ · Immig	-0.0781*	-0.0741*	0.0376*	0.0312*	-0.00703	-0.210*	0.00865
	(0.00477)	(0.00626)	(0.00404)	(0.00249)	(0.00545)	(0.0471)	(0.0213)
CAE$_{15}$ · Immig	-	-0.0855*	-	-0.0681*	-	-	-
		(0.0319)		(0.0169)			
CAE$_{16}$ · Immig	-0.0371*	-0.0878*	0.0129	0.132*	0.0685***	-	0.237*
	(0.00984)	(0.0157)	(0.0144)	(0.00529)	(0.0413)		(0.0364)
CAE$_{17}$ · Immig	0.0464*	0.0286*	0.0498*	0.0978*	0.0911*	-	0.0669*
	(0.00582)	(0.00662)	(0.00400)	(0.00257)	(0.00801)		(0.0230)
CAE$_{18}$ · Immig	0.854*	0.0830*	0.116*	0.651*	-0.0172	0.363*	0.543*
	(0.0292)	(0.0133)	(0.0266)	(0.0246)	(0.0430)	(0.0617)	(0.0438)
CAE$_{19}$ · Immig	-0.00684	0.0244*	0.0370*	0.0810*	-0.0314**	-0.100***	0.0303
	(0.00860)	(0.00815)	(0.00669)	(0.00374)	(0.0136)	(0.0547)	(0.0306)
CAE$_{21}$ · Immig	-	-	-	0.316*	-	-	-
				(0.0412)			

Tenure · Immig	-	-	-	-0.00219* (0.000176)	-	-	0.00558* (0.00137)
Tenure² · Immig	-	-	-	-	-	-	-
$Qual_1$ · Immig	0.160* (0.00801)	-0.104* (0.0115)	0.0602* (0.0105)	0.0707* (0.00455)	-0.0272 (0.0178)	0.167* (0.0487)	0.0953* (0.0276)
$Qual_2$ · Immig	0.193* (0.00755)	-0.0630* (0.00928)	-0.0786* (0.00619)	-0.00553 (0.00374)	-0.223* (0.00929)	0.397* (0.0466)	0.207* (0.0262)
$Qual_3$ · Immig	0.0400* (0.00287)	-0.0456* (0.00397)	-0.0327* (0.00243)	-0.0312* (0.00163)	-0.0854* (0.00450)	-0.0172 (0.0144)	-0.00314 (0.0110)
$Qual_4$ · Immig	0.0587* (0.00325)	-0.0114* (0.00424)	0.00123 (0.00260)	-0.0212* (0.00175)	-0.0434* (0.00476)	0.0110 (0.0175)	0.0283*** (0.0148)
Constant	0.938* (0.00163)	0.899* (0.00418)	0.980* (0.00202)	0.699* (0.00251)	0.796* (0.00388)	0.940* (0.00810)	0.908* (0.00619)
N	2,076,082	352,982	1,271,556	4,597,886	353,939	86,315	131,340

Notes: (i) In parentheses are the robust standard errors.

(ii) *, **, *** denote significant at 1%, 5%, and 10%, respectively.

(iii) "Immig" is the "Immigrant" variable.

(iv) The above results define the optimal number of control variables assuming the exogeneity of the variable of interest.

Source: Quadros de Pessoal. Own elaboration.

Annex H - Selection bias test result

Table 28 - Heckman two-step selection model (2011 - 2020)

(Outcome equation dependent variable: log of the real hourly wage)

(Selection equation dependent variable: “Immigrant”)

	Outcome Equation	Selection Equation
YSM	-0.00196* (0.000221)	21.56
YSM^2	0.000108* (8.42e-06)	8.459
Age	0.00543* (0.000268)	0.0494* (0.00108)
Age^2	-4.97e-05* (3.32e-06)	-0.000792* (1.49e-05)
$Educ_2$	0.0365* (0.000919)	-
$Educ_3$	0.240* (0.00152)	-
$Educ_4$	0.278* (0.00332)	-
$Educ_5$	0.584* (0.00597)	-
$Year_{12}$	-0.0268* (0.00186)	-
$Year_{13}$	-0.0349* (0.00191)	-
$Year_{14}$	-0.0211* (0.00190)	-
$Year_{15}$	-0.0260* (0.00188)	-
$Year_{16}$	-0.0117* (0.00184)	-
$Year_{17}$	0.00859* (0.00178)	-
$Year_{18}$	0.0333* (0.00172)	-
$Year_{19}$	0.0599* (0.00165)	-

$Year_{20}$	0.107* (0.00163)	-
Algarve	0.0227* (0.00166)	-
Centre	-0.0211* (0.00152)	-
Lisbon	0.0298* (0.00126)	-
Alentejo	-0.00885* (0.00198)	-
Azores	-0.0177* (0.00643)	-
Madeira	0.0714* (0.00417)	-
Concentration	-0.0262* (0.000419)	-
Size	0.00787* (0.000251)	-
CAE_1	-0.0448* (0.00212)	-
CAE_2	0.172* (0.00831)	-
CAE_4	0.279* (0.0156)	-
CAE_5	-0.0589* (0.00414)	-
CAE_6	-0.0190* (0.00172)	-
CAE_7	-0.0353* (0.00156)	-
CAE_8	0.00594* (0.00209)	-
CAE_9	-0.0408* (0.00152)	-
CAE_{10}	0.000530 (0.00269)	-
CAE_{11}	0.233* (0.00383)	-

CAE_{12}	-0.0264* (0.00399)	-
CAE_{13}	-0.0155* (0.00263)	-
CAE_{14}	-0.115* (0.00169)	-
CAE_{15}	-0.0804* (0.0130)	-
CAE_{16}	-0.0341* (0.00367)	-
CAE_{17}	-0.116* (0.00192)	-
CAE_{18}	0.643* (0.00360)	-
CAE_{19}	-0.0724* (0.00286)	-
CAE_{21}	0.777* (0.0430)	-
Tenure	0.00948* (0.000231)	-9.747
$Tenure^2$	-4.56e-05* (1.08e-05)	-4.700
$Qual_1$	0.774* (0.00254)	-0.345* (0.00830)
$Qual_2$	0.416* (0.00238)	-0.277* (0.00865)
$Qual_3$	0.0623* (0.00191)	0.0652* (0.00670)
$Qual_4$	0.00920* (0.00210)	0.124* (0.00738)
Brazil	-0.218* (0.00142)	-
CEEC	-0.228* (0.00149)	-
China	-0.232* (0.00257)	-
PALOP	-0.243* (0.00158)	-

SA	-0.211* (0.00196)	-
Others	-0.180* (0.00170)	-
Constant	1.164* (0.00584)	-1.986* (0.0189)
N	8,870,100	
IMR	-0.00476* (0.000736)	

Notes: (i) In parentheses are the robust standard errors.

(ii) *, **, *** denote significant at 1%, 5%, and 10%, respectively.

Source: Quadros de Pessoal. Own elaboration.

FACULDADE DE **ECONOMIA**