ESSAYS ON THE LEGAL AND ECONOMIC INTEGRATION OF REFUGEES

by

HIROMI YUMOTO

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Department of Management

Birmingham Business School

School of Social Science

University of Birmingham

July 2023

Abstract

This thesis consists of three essays on the legal and economic integration of recent refugees in the Netherlands. In Chapter 1, we investigate how assignment location, specifically urban and co-national density, affects citizenship acquisition. Exploiting the quasi-random allocation of refugees within the country, our analysis indicates that refugees in urban areas are more likely to integrate, particularly those with higher education levels. Co-national density is also positively associated with integration, however, its impact diminishes towards the upper end of the sample range. Chapter 2 explores whether the labour market training component (the ONA) enhances refugees' economic integration. We compare the labour market trajectories of refugees required to take the ONA component with those who were not. We find employment probability to be 6.4-11.0 percentage points higher for those who took the ONA but find insignificant differences in hours worked or hourly wages. Chapter 3 focuses on pay disparities, highlighting the considerable wage gaps between refugees and natives, attributing the gaps to firm effects through firm sorting and pay-setting. We show significant pay differentials, with refugee wages approximately half of those of natives. Our analysis also suggests that firm sorting accounts for 14% of the native-refugee pay gap while the

pay-setting effect is essentially zero.

To my inspiring parents.

Declaration

I certify that the thesis presented for examination for the PhD degree of the University of Birmingham is solely my own work other than where collaboration is specifically indicated and that I have fully acknowledged all sources of information that have been used in the thesis. I declare that this work has not been submitted for a degree at any other university or institution. The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This doctoral thesis may not be reproduced without the author's written consent. I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party, and contains fewer than 80,000 words, including appendix, bibliography, footnotes, tables and equations.

A version of Chapter 2 was submitted to the Journal of Labor Economics (The University of Chicago Press) in June 2023.

Hiromi Yumoto

July 2023

Acknowledgements

This thesis has been a long journey filled with challenges, which involves many people to who I am indebted. I offer my sincere appreciation and deepest respect. Thank you for being an integral part of my story and for helping me reach this momentous milestone in my academic career.

I am profoundly indebted to my exceptional supervisors, Professors Matthew Cole, Liza Jabbour and Ceren Ozgen. Their brilliant guidance, wisdom, and dedication to my research have been invaluable. They challenged me with thought-provoking questions and provided constructive feedback that shaped my work and elevated its quality to meet high research standards. Moreover, their support extended beyond academics, as they also showed genuine concern for my well-being throughout my doctoral journey at the University of Birmingham.

I would also like to gratefully acknowledge the financial support from the Joint Japan/World Bank Graduate Scholarship Program (JJ/WBGSP). Without their generous financial support, this PhD journey would not have been possible.

Furthermore, I would like to express my gratitude to the Departments of Management and Economics at the University of Birmingham for providing an intellectually stimulating and supportive research environment. Additionally, I am thankful to *Dienst Uitvoering Onderwijs* (DUO) for their assistance in providing necessary background information about the Dutch civic integration exams for refugees, which enriched my research.

To my dear friends, whose unwavering moral support, attentive listening, and a healthy dose of humour have been invaluable, I offer my heartfelt thanks. Your encouragement, advice, and wit have kept me grounded and motivated throughout this challenging process, and I am profoundly grateful for your unwavering friendship.

At last, but certainly not least, I would like to express my deepest gratitude to my family and my partner, Andy. You have consistently been my rock and foundation, providing me with unconditional love and support, and making tremendous sacrifices on my behalf. Your presence has sustained me through the good and bad times in my challenging academic life, and I am profoundly grateful for everything you have done for me.

Contents

List of Tables

List of Figures

Introduction

In recent years mass migration has been observed around the world as globalisation has deepened further. While immigrants surged from 84 million to 271 million in the last four decades, forcibly displaced persons reached 108 million by the end of 2022 as a result of human rights violations, persecution, and armed conflicts (IOM, 2019; UNHCR, 2023). In particular, the conflicts in Syria since 2014 and the recent Russian invasion of Ukraine forced an unprecedented number of people to head towards Europe for asylum. The countries that face a sudden refugee influx attempt to promote their integration into society not only for humanitarian purposes but also to minimise their economic dependency on public welfare systems, to reduce any threat to social cohesion and to limit economic segregation within the host countries (Enchautegui and Giannarelli, 2015; Adamson, 2020; Foged and Werf, 2022). Despite the potentially significant costs associated with a lack of refugee integration, there is surprisingly little evidence on how well refugees adjust to new environments and what factors help or hinder their social and economic participation. The lack of reliable information on integration obstacles and outcomes means policymakers have little guidance when designing asylum and immigrant policies.

Varela et al. (2020) claim that integration is achieved in multiple dimensions: legal, economic, and cultural. While cultural integration relates to respect for the cultural and religious diversity, legal and economic integration imply equal opportunities and rights in political and economic activities, including access to the labour market, health services and voting rights. This thesis focuses on economic and legal integration.

Full legal integration (naturalisation) results in the acquisition of citizenship, providing refugees with a full range of rights, including the ability for their voices to be heard through the electoral process (Huddleston and Vink, 2015). Indeed, in the absence of legal integration, refugees' inability to vote may deepen the problems they face due to their limited visibility to local and national governments. This lack of feedback from refugee communities combined with limited academic evidence results in a dearth of guidance for policymakers on how to encourage social, economic or legal participation. It has therefore been claimed that naturalisation is the most advanced outcome of the integration processes in the host countries (Hoon et al., 2019).

The other aspect of integration this thesis focuses on, economic integration, is vital to prevent economic marginalisation. Economic integration, which implies refugees have equal opportunities in the labour market, can be reached through labour market participation, securing a job consistent with refugees' skills or qualifications, and being paid fairly and equally.

As the economic integration of refugees deepens, it is expected that gaps between immigrants and native-born workers in the labour market will converge. The process of reducing these gaps, especially wage gaps, is defined as the economic assimilation of immigrants (Borjas, 2015).

Immigrants will often face challenges in achieving legal and economic integration, often due to limited proficiency in the local language, difficulties in transferring qualifications obtained outside the host country, and a lack of familiarity with local norms. Refugees experience additional disadvantages throughout the integration process. The difficulties experienced by refugees have been well-documented and stem from factors such as disrupted education or working tenures and poor health, potentially as a consequence of trauma and discrimination (see, for example, Fasani et al. (2022)). These difficulties typically manifest themselves in terms of low pay, a greater likelihood of being over-educated, and a lower likelihood of securing employment. The existing literature, however, tends to overlook the experiences of refugees despite these additional challenges (Bevelander, 2020; Brell et al., 2020).

This series of essays aims to assess the determinants of, and obstacles to, the economic and legal integration of refugees in the Netherlands. Like much of Europe, the Netherlands received substantial flows of refugees since 2014, with refugees per capita being consistent with the EU as a whole. It, therefore, offers an ideal setting to analyse refugee integration in the European context (for more detail, please see Chapter 1). Using Dutch administrative microdata including the asylum cohort study *Asielcohort*, wage data (*SPOLISBUS)* and firm census data (*Production Statistice*) for the period between 2014 and 2022, this study assesses the factors that help and hinder the legal and economic integration of refugees.

Chapter 1 investigates the effect of locational factors on acquiring citizenship. In assessing the role played by assignment location it specifically focuses on the effects of urban and co-national density, exploiting the quasi-random allocation of refugees within the Netherlands. It uses two measures of legal integration, first whether or not an individual passed the civic integration exam and second, whether full naturalisation occurred. Based on OLS and IV analysis it is found that refugees allocated to urban areas are more likely to integrate, while the duration of employment in urban areas also increases the likelihood of integration. The effects of urbanisation are greatest for those with high levels of education. Co-national density is also positively associated with integration, however, like urbanisation, its impacts become negative towards the upper end of the sample range.

Chapter 2 explores whether a civic integration component dedicated to labour market training (the ONA) boosts refugees' economic integration. To complete the ONA component, refugees are required to research the types of jobs they are interested in, identify the qualifications or skills for these specific jobs, prepare a CV, and sit mock job interviews. Using linked employer-employee data from 2014 to 2021 for the Netherlands, Chapter 2 compares the labour market trajectories of refugees required to take the ONA component with those who were not. Employment probability is found to be 6.4-11.0 percentage points higher for those who took the ONA but no significant differences in hours worked or hourly wages are found. In addition, the ONA does not improve the persistent labour market disadvantages faced by females and Syrians. Finally., this chapter also explores the mechanisms and firm effects through which the ONA operates.

Chapter 3 explores the role of firms in refugee-native pay disparities. Drawing on the existing literature on pay gaps, this chapter highlight firm effects on total pay disparities via two complementary channels: firm sorting; and pay-setting. Firm sorting focuses on differences in the fractions of natives and refugees employed in a firm with specific pay premiums; while the second channel emphasises the differences in wage-setting within firms. The results show significant pay differentials between refugees and natives, with refugee wages approximately half of those of natives. A substantial mismatch between refugee ability and firm-specific pay premiums among refugees is also found. The analysis also suggests that individual characteristics are the main drivers of pay gaps *within* and *between* groups, though firm sorting accounts for 14% of the native-refugee pay gap. The pay-setting effect is close to zero.

The remainder of this doctoral thesis is organised as follows. The first essay titled "Invisible Immigrants: Urbanisation, Co-national Density and the Legal Integration of Refugees" is presented in Chapter 1. The second essay on "Local Labour Market Training and Civic Integration Exams: Examining the Employment Outcomes" is presented in Chapter 2, followed by the third essay, "Firm Sorting and the Economic Assimilation of Refugees" in Chapter 3. The thesis concludes with some final remarks, including potential pathways for future research.

Chapter 1

Invisible Immigrants: Urbanisation, Co-National Density and the Legal Integration of Refugee

Abstract

We use rich micro-data on refugees who arrived in the Netherlands between 2014 and 2020 to analyse the factors that influenced their decision to legally integrate. We focus on the role played by assignment location and specifically the effects of urban and co-national density. Our analysis benefits from the quasi-random allocation of refugees within the Netherlands and uses two measures of legal integration, first whether or not an individual passed the civic integration exam and second, whether full naturalisation occurred. We find that refugees allocated to urban areas are more likely to integrate while the duration of employment in an urban area also increases the likelihood of integration. The effects of urbanisation are greatest for those with high levels of education. Co-national density is also positively associated with integration, however, like urbanisation, its effects become negative towards the upper end of the sample range. We also find that mental or physical health difficulties, a lengthy asylum application process, and the duration of time on social benefits reduce the likelihood of integration. In contrast, both forms of legal integration are positively associated with being in employment and being educated within the Netherlands.

Keywords: Refugees, legal integration, naturalisation, urbanisation

JEL classification codes: J15; J18; O52; R19

1.1 Introduction

Recent decades have seen a significant increase in the movement of people across the world. Globally, the number of immigrants increased from 84 million to 271 between 1970 and 2019, while the number of forcibly displaced persons reached 80 million in 2020 as a result of human rights violations, persecutions, and armed conflicts (IOM, 2019; UNHCR, 2021). The number of refugees seeking protection in Europe alone reached 6 million in 2019, in part caused by large numbers of people fleeing the conflict in Syria in 2014-15 (*Ibid.*). By November 2022, due to the war in Ukraine 7.8 million refugees were recorded across Europe.[1]

While refugees have often faced physical or psychological trauma in their countries of origin, this trauma may be compounded by a number of challenges upon arrival in host countries. Language difficulties, broken education and cultural differences can result in a lack of opportunities, potentially fuelled by discrimination. In this context, the successful integration of refugees - whether social, economic or legal - has become a key challenge within host countries. It is clear that a lack of integration, perhaps coupled with a dependency on welfare systems, may threaten social cohesion and worsen economic segregation in host societies (Foged et al., 2022; Adamson, 2020; Enchautegui and Giannarelli, 2015). In turn, this dependency feeds the politicisation of immigration and can cause a divide between policy makers and immigrant communities.

[1] https://data.unhcr.org/en/situations/ukraine

Despite the importance of refugee integration and the problems that may arise where it is lacking, we currently have little knowledge or understanding of the factors that may encourage or hinder it. Full legal integration (naturalisation) results in the acquisition of citizenship, providing refugees with a full range of rights including the ability for their voices to be heard through the electoral process (Huddleston and Vink, 2015). Indeed, in the absence of legal integration, refugees' inability to vote may deepen the problems they face due to their limited visibility to local and national governments. This lack of feedback from refugee communities combined with limited academic evidence results in a dearth of guidance for policymakers on how to encourage social, economic or legal participation. Limited academic evidence is particularly problematic given the complexity of refugee resettlement programmes, their costly nature and the challenges associated with receiving the support of host communities. The existing literature tends to treat naturalisation as a stepping stone to economic integration rather than establishing the factors that select refugees into obtaining citizenship. The only comprehensive empirical study to date is Mossaad et al. (2018) who analyse refugees within the United States (US) and find that the characteristics of arrival cities may influence legal integration (naturalisation) together with individual characteristics such as country of origin and human capital. Very recently, Eckert et al. (2022) confirm the importance of arrival cities in the case of Denmark, yet only for refugee wage growth. However, the absence of broader evidence for Europe, where refugee patterns are quite different to those in the US, means policymakers have little guidance on how to design effective asylum and integration programmes for more than 10 million refugees in Europe.

In this context, we provide the first analysis of the recent influx of refugees into Europe and identify the factors influencing their *naturalisation* in the case of the Netherlands, a country that received almost 45,000 asylum applications in 2015 alone. We are unaware of any existing studies that examine the naturalisation of the recent influx of refugees into Europe. The European evidence of refugee integration that does exist typically focuses on a handful of countries (e.g. Denmark, Norway, Switzerland) and examines the refugees that arrived prior to 2014 and/or focuses only on their integration into labour markets. In particular, studies of the economic integration of refugees in the recent literature typically focus on i) asylum application processing times (Gathmann and Keller (2018); Hainmueller et al. (2016)); ii) linguistic training (Foged et al. (2022)); iii) locational characteristics, e.g. initial assignment location (city-size effects in Eckert et al. (2022)) or local co-national networks (Martén et al. (2019); Stips and Kis-Katos (2020); Godøy (2017)).[2]

While these studies are clearly useful, participation in labour markets alone represents significantly less integration than full citizenship, meaning that the factors influencing the two

[2] For example, Fasani et al. (2022) examine refugees into European Union (EU) member states prior to 2013 and find their labour market outcomes to be inferior to those of other immigrants. Similarly, Hainmueller et al. (2019) investigate the naturalisation process of refugees who entered Switzerland between 1970 and 2003 and show the benefits of such naturalisation in terms of subsequent average earnings. Foged et al. (2022) assess refugee cohorts who arrived in Denmark around 1999 and demonstrate the potential benefits of language training to future earnings. Kone et al. (2019) and Stips and Kis-Katos (2020) also include recent refugees or asylum seekers in Europe in their samples but again focus on their labour market performance which is shown to be inferior to other immigrants but positively influenced by co-national networks.

could be quite different. Similarly, the scale and ethnic composition of recent refugees into Europe is unlike that in previous years which again could influence naturalisation outcomes. Most importantly, economic integration is only part of the visibility refugees receive from central governments as without full naturalisation they have no voting power at national elections. Therefore, as far as we are aware, no existing studies analyse the factors that influence the legal integration of refugees into Europe.

Our study utilises highly detailed administrative data on the universe of refugees who arrived in the Netherlands between 2014 and 2020 to explore the factors that influenced their integration. To do so we distinguish between two forms of integration. The first captures whether or not an individual has passed the integration exam, a necessary condition in order to apply for permanent residency or full naturalisation. The second is whether or not full naturalisation was attained.[3] We benefit from a quasi-experimental setting as a result of a dispersion policy in which the Dutch authorities allocate refugees to reception centres and ultimately municipalities on a quasi-random basis. In addition to the detailed economic and social characteristics of refugees we are therefore also able to examine the roles played by fine-grained locational characteristics. This allows us to focus on the effect of two broad channels emanating from the recent urban economics literature. First, we explore the *density* channel. If learning, knowledge diffusion, participation in labour markets and returns to employment are all greater in urban areas (Duranton and Puga (2004); De La Roca and

[3] It is necessary to have resided in the Netherlands for at least 5 years to apply for naturalisation although the integration exam can be taken prior to this.

Puga (2017)) then we might expect urban density, and the duration of urban employment, to influence naturalisation. Second, we consider the *co-national* channel. If the density of co-nationals is high, refugees may benefit from greater transmission and acquisition of information through bonding social capital (Patacchini and Zenou (2012); Klaesson et al. (2021)), perhaps positively influencing the likelihood of integration. However, a point may be reached after which co-national density negatively impacts bridging social capital - capital generated through interactions between different groups of people - which may reduce the likelihood of integration (Putnam, 2007). Within this co-national channel we also consider the distinction between co-national density and the density of *employed* co-nationals (Klaesson et al. (2021); Stips and Kis-Katos (2020)). We also examine whether the effects of urbanisation on legal integration differ with the education level of refugees which may arise if more able individuals enjoy greater learning advantages from big cities (De La Roca and Puga (2017)).

To briefly summarise our results, we find that urbanisation matters for legal integration. Refugees located in densely populated areas are more likely to legally integrate unless density reaches a very high level, while the duration of employment in urban areas also increases the likelihood of integrating. We additionally find some evidence to suggest that these effects are greatest for those with higher levels of education. The co-national channel is also found to be important and positively influences the likelihood of integrating but again becomes negative at the upper end of the sample range. A number of other factors are found to reduce the likelihood of legal integration, including mental or physical health difficulties, the duration of time on social benefits and a lengthy asylum application process.

The remainder of the paper proceeds as follows. After reviewing the theoretical background and the empirical literature on refugee integration in Section 2, Section 3 overviews the Dutch asylum system and provides related statistics. Section 4 describes the data and the sample used for the estimation. Section 5 discusses the methodological approach while Section 6 presents the findings, including robustness checks. The last section concludes.

1.2 Determinants of refugee integration

The integration of foreign-born populations into host countries has several dimensions including social, economic, cultural, legal and political (Varela et al., 2020). The focus of this paper is on *legal* integration, the strongest form of which is citizenship or naturalisation. The acquisition of citizenship provides immigrants with equal rights and opportunities to citizens of the host country including access to the labour market, health services, and voting rights. Naturalisation is perceived as the best guarantee for the immigrants' rights, equal treatment, political participation, integration into society (Huddleston and Vink, 2015; Engdahl et al., 2020) and integration into labour markets (Gathmann and Keller, 2018). Gaining citizenship usually implies an intention to reside in the host country in the long term, even permanently.[4]

[4] Nevertheless, some people apply for citizenship in the EU to move into other European countries. In this case, naturalisation would work as 'a ticket to mobility' (Hoon et al., 2019).

Without opportunities to express their voices to policymakers through formal channels, the demands and problems of immigrants in general, and refugees in particular, are likely be overlooked. In other words, immigrants, including refugees, without full legal integration are neither visible to policymakers nor politically engaged with the issues of their host countries.

There is a growing body of literature that analyses the relationship between the acquisition of citizenship and the integration of immigrants. This literature is mostly focused on the implication of naturalisation on the labour market outcomes of immigrants (Gathmann and Keller, 2018; Hainmueller et al., 2019; Govind, 2021) and is anchored in the political debate on whether naturalisation should be used as a 'catalyst' for further integration, by providing immigrants with the means and incentives to invest in integration, or as a reward for immigrants who have achieved full integration and assimilation into host countries (Hainmueller et al., 2017). While this literature acknowledges the self-selection of immigrants into applying for and acquiring citizenship, there is limited evidence on the characteristics of naturalised immigrants or on the determinants of the naturalisation process.[5] Moreover, with the exception of Mossaad et al. (2018) who analyse the naturalisation of refugees, the existing literature primarily focuses on immigrants yet the distinction between immigrants and refugees might be particularly important because of the differences in the arrival conditions of these

[5] A notable exception is the work by Bevelander and Veenman (2006) who analyse the link between naturalisation and socioeconomic integration in the Netherlands.

two groups.[6] In contrast, the vast majority of the literature on refugees integration focuses on the economic dimension and particularly on labour market outcomes. We here draw on these bodies of literature to identify the main determinants of the legal integration of refugees.

Mossaad et al. (2018) present a comprehensive analysis of the determinants of refugee naturalisation in the US and show that naturalisation outcomes differ substantially with refugees' individual characteristics. For instance, well-educated refugees naturalise at higher rates with university graduates being more likely to gain citizenship than refugees with no education. Country of origin is also a significant predictor of naturalisation. On the other hand, there is no significant effect associated with family characteristics such as family size or position in the family. These findings are in line with the extant literature (e.g.Portes and Curtis (1987); Yang (1994)) that also indicates that human and social capital boost the probability of naturalisation.

Other studies indicate that the economic integration of foreign-born workers is significantly affected by individual characteristics, such as the duration of residence in host coun-

[6] Compared to other types of immigrants, the poor labour market performance of refugees is common in advanced countries (De Vroome and Van Tubergen, 2010; Kone et al., 2019). Some empirical studies, including in Europe such as Fasani et al. (2022), show large initial labour market disadvantages of refugees which cannot be fully explained by observable characteristics such as age, gender and education. These disadvantages persist for long after migration. In the same way as other workers, labour market performance also depends on gender. As Bevelander (2020) finds, male refugees close the gap with the employment rate of immigrants more quickly than females. The gaps among female refugees and immigrants persist even two decades after arrival. Such disadvantages can be explained by refugee-specific difficulties, such as mental health problems caused by traumatic experiences in their countries of origin, the long legal process for asylum claims, and the restricted access to labour markets (Foged et al., 2022; Brell et al., 2020).

tries and country of origin (e.g. Brell et al. (2020); De Vroome and Van Tubergen (2010). In particular, educational attainments are key to integrating refugees into the local labour market (Card, 2009; Yao and van Ours, 2015). However, human capital earned/recognised in host countries can be more effective than education prior to arrival (Foged et al., 2022; Brücker et al., 2021).

Recent papers within urban economics suggest that urbanisation might influence the likelihood of integration. While not themselves addressing immigrant integration, De La Roca and Puga (2017) argue that the benefits of agglomeration economies are greatest in urban areas perhaps indicating that the ability and desire to integrate might be enhanced by urban density. Similarly, Eckert et al. (2022) show that between 1986 and 1998, those refugees placed in Copenhagen exhibited 35% faster wage growth with each additional year of experience compared to those placed in other areas in Denmark.

Studies also suggest that refugees in communities with dense co-national populations would benefit from social networks that may provide access to information, financial support and ethnic amenities (McKenzie and Rapoport, 2010; Ozgen et al., 2013). The presence of co-nationals or a foreign-born population in the receiving community may also assist economic integration although the evidence is mixed (Braun and Dwenger, 2020; Martén et al., 2019; Godøy, 2017). Klaesson et al. (2021) and Patacchini and Zenou (2012) provide empirical evidence that *employed* ethnic peers are particularly important in supporting new immigrants to obtain their first job while Stips and Kis-Katos (2020) point out a negative link between the employment probability of refugees and the number of non-employed co-nationals. Klaesson et al. (2021) also emphasise the distinction between co-national shares *per se* and *employed* co-national shares.

Mossaad et al. (2018) demonstrate that the characteristics of arrival cities can have a significant impact on naturalisation decisions. For instance, they find that a higher unemployment rate in the arrival city reduces the naturalisation probability, whereas a high share of co-nationals is associated with a higher likelihood of becoming an American citizen. It is worth noting that the links between regional characteristics and integration outcomes may present endogeneity problems. The presence of co-nationals in the same city is expected to facilitate integration through larger co-national networks and ethnic amenities. Therefore a quasi-experimental setting like ours is likely to better address these concerns.

Finally, the literature points out that the length of asylum applications adversely affects the integration of refugees via deteriorated human capital and difficulties associated with mental health (Brell et al., 2020). Hainmueller et al. (2016) find that one additional year of waiting for the asylum decision reduces the subsequent employment rate by 4 to 5 percentage points. They also suggest the channels through which the adverse impacts of asylum application periods operate: skill atrophy; and physiological discouragement. While skill atrophy refers to a depreciation of skills due to an absence from work during the asylum application, the latter refers to mental health issues caused by their vulnerable status under the threat of deportation if their asylum claims are declined.

1.3 Refugees in the Dutch context

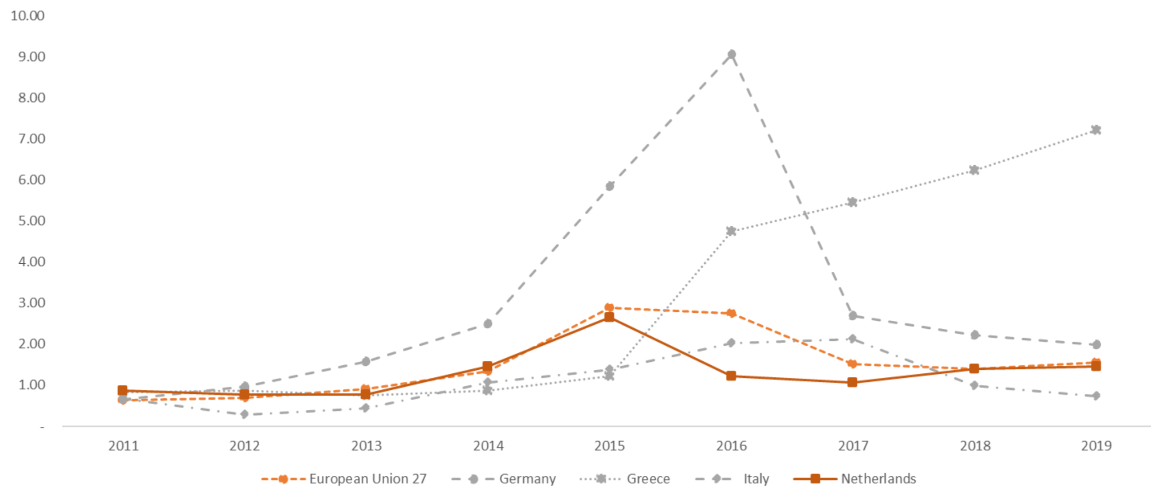
1.3.1 Refugees in numbers

Figure 1.1 provides the trends in asylum applications for the Netherlands, a selection of EU countries, and the EU as a whole, with each expressed relative to 2011 and in per capita terms. As Figure 1.1 indicates, by 2015 asylum applications in the Netherlands (which equalled 45,000 in number) were approximately 3 times their 2011 level, while the EU average was closer to 4.5 times the 2011 level. In large part these increases were a result of the violence in Syria which forced a large number of people to flee. In Germany, asylum applications peaked in 2016 at approximately 14 times their 2011 level while Greece, unlike the rest of the EU, did not experience a reduction in asylum applications after 2015/16 with numbers still increasing in 2019.[7]

Among the recent refugees in the Netherlands, the Central Bureau of Statistics (CBS) reveals that only a quarter of those of working-age who were granted an asylum residence

[7] In terms of the impact on the labour force, the OECD (2019c) reports that refugees in the Netherlands constitute a 0.2 per cent increase in the labour force, compared to 0.8 per cent in Germany, 0.4-0.6 per cent in Greece and a smaller increase of 0.1-0.15 per cent in other European countries such as France.

Figure 1.1: Asylum applications per capita in Europe and the Netherlands



Note: The number of applications per 1,000 population. Source: Eurostat (2021); World Bank (2021)

permit in 2014 had a job three and a half years later (CBS, 2019). In terms of employment contracts, the vast majority of the employed refugees worked as part-time and/or temporary workers (81 and 89 per cent, respectively). Compared to the overall foreign-born employment rate (64.9 per cent according to OECD (2019b)), asylum residence holders, who are able to work without restrictions, performed notably worse. As a result of poor economic participation, approximately 90 per cent of working-aged refugees were receiving social benefits one and a half years after obtaining their residency permit.

CBS (2019) show that nearly 40 per cent of the 2014 refugee cohort enrolled in Dutch education programmes in 2018 with many joining secondary or vocational training rather than academic education. CBS point out that this is probably due to a lack of understanding in the areas of the Dutch language and arithmetic, which are necessary for higher education.

De Vroome and Van Tubergen (2010) and De Vroome et al. (2011) empirically analyse the integration of refugees in the Netherlands, using survey data collected in 2003. Even though the composition of the refugees' countries of origin[8] differ from that of recent refugees, they found that 74.4 per cent of the sample held Dutch nationality after spending 9.5 years in the Netherlands on average (De Vroome et al., 2011). Ersanilli (2014) claims that the naturalisation rate among all immigrants in the Netherlands is generally high compared to other European countries. Hoon et al. (2019) also point out that over 80 per cent of the immigrants who arrived in the Netherlands from 1995 to 1999 had Dutch citizenship by the end of 2015. On the other hand, the labour market participation rate among the sample, aged between 18 and 64, was only approximately 20 per cent (De Vroome and Van Tubergen, 2010).

1.3.2 Asylum policy and application procedures

The current legal framework for asylum in the Netherlands is called the Alien Act of 2000 (*Vreemdelingenwet 2000: Vw 2000*) and The Civic Integration Act (*Wet inburgering: Wi*)[9]. Following a successful asylum application, individuals are granted an asylum residence permit. After five years of legal residency they can apply for Dutch citizenship or a Permanent Residence Permit (PRP).

[8] The top five countries of origin in their sample are Afghanistan, Iraq, Somalia, Iran and the former Yugoslavia. On the other hand, the recent refugees are mainly from Syria and Eritrea.

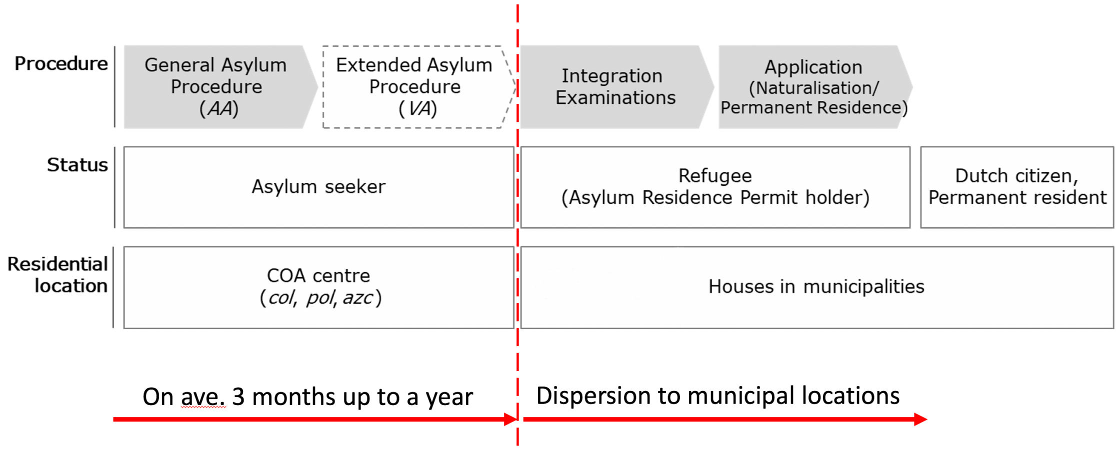
[9] The first version, which was in force in 2007, followed by the amended version in 2013 and 2021.

Those who look for asylum in the Netherlands have to follow the application procedures for an asylum residence permit as shown in Figure 1.2. After arrival asylum seekers go through the General Asylum Procedure (*AA: Algemene Asielprocedure*) at refugee centres run by the Central Agency for the Reception of Asylum Seekers (*Centraal Orgaan opvang asielzoekers*: *COA*). At the end of this process, the application will either be rejected or an asylum residence permit will be issued, sometimes after an Extended Asylum Procedure (*VA: Verlengde Asielprocedure*).[10]

During the period of asylum application and after being granted asylum, refugees receive shelter, meals, access to the necessary medical care at the COA centres, and are relocated as the asylum claim proceeds. At the onset of the asylum claim, they stay in one of the central reception centres (*centrale ontvangstlocatie: COL*) in Ter Apel, Budel-Cranendonck and Veenhuizen. When the asylum application begins, asylum seekers will move to a Process reception centre (*procesopvanglocatie: POL*). Once asylum seekers have obtained an asylum residence permit and become refugees, they move to an asylum seekers reception centre (*Asielzoekerscentrum: AZC*) located across the Netherlands to wait for a house in a municipality. When refugees move to *AZC* or housing after the asylum application procedure they

[10] According to CBS, on average asylum seekers waited for 109 days to obtain the asylum residence permit during the period between 2014 and the first half of 2018 (CBS, 2019). Compared to other EU countries, the application duration in the Netherlands is substantially shorter. Bertoli et al. (2020) estimate the average application processing time in Germany, which accepted over 1 million refugees in 2014, was approximately 16 months while other European countries took approximately 9 months.

Figure 1.2: Standard procedures for asylum and naturalisation



Note: If an asylum claim under the General Asylum Procedure *AA* is accepted, the individual does not go through the Extended Asylum Procedure *VA*.
The information in this figure is from the Immigration and Naturalisation Service and the Central Agency for the Reception of Asylum Seekers (*Centraal Orgaan opvang asielzoekers*) and IND (2019) and is visualised by the authors.

are allocated randomly (Tolsma et al., 2021). This issue is returned to in Section 4.

The rights of asylum status holders depend on the type of permit that they hold. During the application process asylum seekers are allowed to work only after six months with a maximum duration of employment (24 out of 52 weeks) while refugees with asylum status can work without a working permit or any constraints. Also, they are allowed to choose their residences only after gaining the permit and being financially independent and to travel abroad with a valid passport.

In order for refugees to become eligible for Dutch citizenship a number of conditions must be met. These include being 18 years old or above, having lived in the Netherlands for at least five years with a valid residence permit, and being sufficiently integrated into Dutch society (IND, 2019). To prove a sufficient level of integration, they have to pass the civic integration examination or meet the exemption requirements such as obtaining a diploma taught in Dutch or providing evidence of medical difficulties.[11]

1.4 Data and Sample

1.4.1 Data

This study utilises Dutch administrative data of refugee cohorts entitled *Asielcohort*, which is provided by the Dutch Central Bureau of Statistics (CBS). It covers all refugees and asylum seekers who arrived in the COA centres or received an asylum residence permit for the period 2014 to 2020. The total number of individuals included in the data is over 260,000. The dataset contains very detailed personal information such as country of origin, the location of the COA centre they registered in, current location, education obtained in the Netherlands, health conditions, economic status and earnings, and allows us to construct the majority of the variables used in this analysis.

The sample for the estimation consists of all post-2013 refugees in the Netherlands who

[11] Before the enforcement of the Civic Integration Law 2007, immigrants had to participate in civic integration courses instead of passing the exam(Ersanilli, 2014). The amended of the law in 2013, which places the responsibility with migrants themselves rather than municipalities, requires immigrants to pass the integration exam (Henskens and Dijsselbloem, 2020).

meet the basic requirements for eligibility to sit the integration exam and obtained an asylum residence permit in the period. In other words, the sample contains all those who meet the following criteria: (1) aged 18 and over (i.e. eligible to request that they sit the integration exam)[12]; (2) those who arrived in 2014 or later and obtained an asylum residence permit;[13] (3) those who stayed in the same municipality after leaving the COA centre.[14] After constructing our sample using these criteria we are left with 48,001 individuals.

We complement this dataset with municipality-level information including unemployment rate and population density from StatLine, matching to *Asielcohort* using municipality codes. In addition, the *BRP* (Personal Record Database), which records the universe of residents in the Netherlands, is also used to calculate the share of co-national population as well as to identify refugees who have acquired Dutch citizenship. Furthermore, the wage data *SPOLIS*, which contains labour market outcomes of all employed individuals in the Netherlands, is used to compute hours worked and whether or not an individual is in employment.

---

12 Individuals can be granted exemption from the integration exam on the basis of special individual circumstances or medical circumstances. To be granted such exemption individuals have to formally apply for it which, in some cases, incurs a charge. Since such individuals have to actively seek out an exemption, we include them within our sample alongside those who pass the exam as both groups are actively choosing this necessary step within the legal integration process. In the robustness section we omit those who are granted exemption from the exam from our sample. For simplicity the remainder of the paper refers to those who passed or were exempt from the exam as the 'passed exam' sample. The proportion of refugees who fail the exam is very low (less than 4.5 per cent). Their inclusion in the sample has no discernible effect on our results and so we have chosen to omit them from our analysis.

13 We exclude those who arrived prior to 2014 but who were granted an asylum residence permit in 2014 or later. Such individuals have been delayed in the system, perhaps due to the complexity of their cases, and hence are unlikely to be representative of refugees more generally.

14 Those who moved out from the initially allocated municipality are removed from our sample because relocation would cause endogeneity concerns. These individuals constitute 12.8 per cent of the sample. Individuals who passed away or who left the Netherlands are also excluded

1.4.2 Our sample's characteristics

As Table 1.1 indicates, 65.8 per cent passed or were exempted from the integration exam while 23.8 per cent received Dutch citizenship by 2020. The reason for the notable gap in integration rates between the exam outcomes and the naturalisation is that naturalisation requires residence for at least five years. In contrast, the exam can be taken prior to completing five years of residency. Those who passed or were exempted from the exam cannot therefore apply for citizenship until they meet the five-year residency condition.

Table 1.1 also provides summary statistics for the sub-sample of refugees who did not pass the integration exam in column (2), compared with those who did in column (3). It indicates that the integrated group is slightly older, stayed in the Netherlands longer, and waited for asylum claims for a shorter length of time. This is in line with expectations as the average time elapsed before taking the integration exam since entry to the Netherlands is around 4 years. Indeed, those who passed the exam resided in the country for 5.1 years and waited for the asylum decision for about 4.2 months while the average residence for those who did not pass the exam is less than 4 years and the asylum decision took for 5.1 months.

A contrast between these 2 groups is with respect to their educational attainment in the Netherlands after arrival. Those who passed the exam were more likely to have no education in the Netherlands (82.6 per cent compared to 73.3 per cent) and specifically less likely to be enrolled in low or middle education (15.0 per cent compared to 25.3 per cent). This is likely to reflect the fact that refugees who passed the exam had higher levels of education from their countries of origin than those who did not pass the exam. Unfortunately we do not have information on education in the country of origin. Our data also show that refugees of Syrian origin constitute by far the largest group within those who passed the integration exam. The detailed variable descriptions are shown in Tables A.1-A.3 in the Appendix.

Table 1.1 additionally presents economic conditions before the integration exam or by 2020 for those who did not pass the exam. It is notable that those who passed the exam received social benefits for a greater duration (as a share of their total time in the Netherlands) than those who did not pass the exam (71.4 per cent compared with 54.1 per cent) but a higher share are also in employment (27.7 per cent compared with 21.3 per cent). Finally, Table 1.1 reveals little difference between the two sub-samples in terms of the characteristics of municipalities.

Table A.4 in the Appendix repeats Table 1.1 and focuses on the cohort of refugees who arrived in 2014 since for this cohort alone we are able to compare those who passed the exam with those who fully naturalised.[15] Table 2 therefore compares the characteristics of the full 2014 cohort sample, those who did not integrate (i.e. didn't pass the exam and hence didn't

[15] Those who arrived after 2014 were not able to naturalise within our sample period due to the five-year residency requirement.

Table 1.1: Summary Statistics (Cohort 2014-2020)

	(1) Full sample		(2) Not integrated		(3) Passed exam	
	Mean or %	SD	Mean or %	SD	Mean or %	SD
Passed exam/Exempted	65.81%					
Naturalised	23.77%					
Years since migration	4.70	1.04	3.94	1.11	5.09	0.75
Arrival year						
2014	24.29%		8.01%		32.75%	
2015	47.62%		33.65%		54.88%	
2016	11.70%		17.35%		8.77%	
2017	13.10%		31.56%		3.53%	
2018	3.27%		9.44%		0.07%	
Asylum application months	4.52	5.27	5.13	7.07	4.20	3.98
Age at arrival	30.73	11.62	30.15	14.09	31.02	10.08
Female	33.71%		40.97		29.94%	
Education in the Netherlands						
No education	79.45%		73.30%		82.64%	
Low or middle	18.50%		25.26%		15.00%	
High	2.05%		1.44%		2.36%	
Country of origin						
Syria	60.21%		52.65%		64.14%	
Ethiopia	10.78%		9.19%		11.61%	
Eritrea	9.89%		9.75%		9.96%	
Iraq	3.45%		5.05%		2.61%	
Iran	3.20%		5.54%		1.99%	
Stateless	9.87%		6.48%		11.62%	
Economic conditions						
Duration of social benefits	65.47%	25.99%	54.07%	32.53%	71.39%	19.33%
Currently employed	25.55%		21.34%		27.74%	
Hours worked in urban	0.23	0.65	0.28	0.77	0.21	0.59
Hours worked in rural	0.12	0.47	0.13	0.51	0.12	0.45
Municipal characteristics						
Unemployment rate	3.85%	0.98%	3.94%	0.90%	3.81%	1.02%
Population density	1.85	1.81	1.97	1.88	1.78	1.77
Co-national density	0.04	0.08	0.04	0.08	0.04	0.08
Employed Co-nat. density	0.01	0.02	0.01	0.02	0.01	0.02
Co-national stock	5.28	12.78	5.74	13.35	5.04	12.47
Health conditions						
Mental health problems	4.94%		4.83%		5.00%	
Chronic physical problems	1.30%		2.27%		0.80%	
Observations	48,001		16,412		31,589	

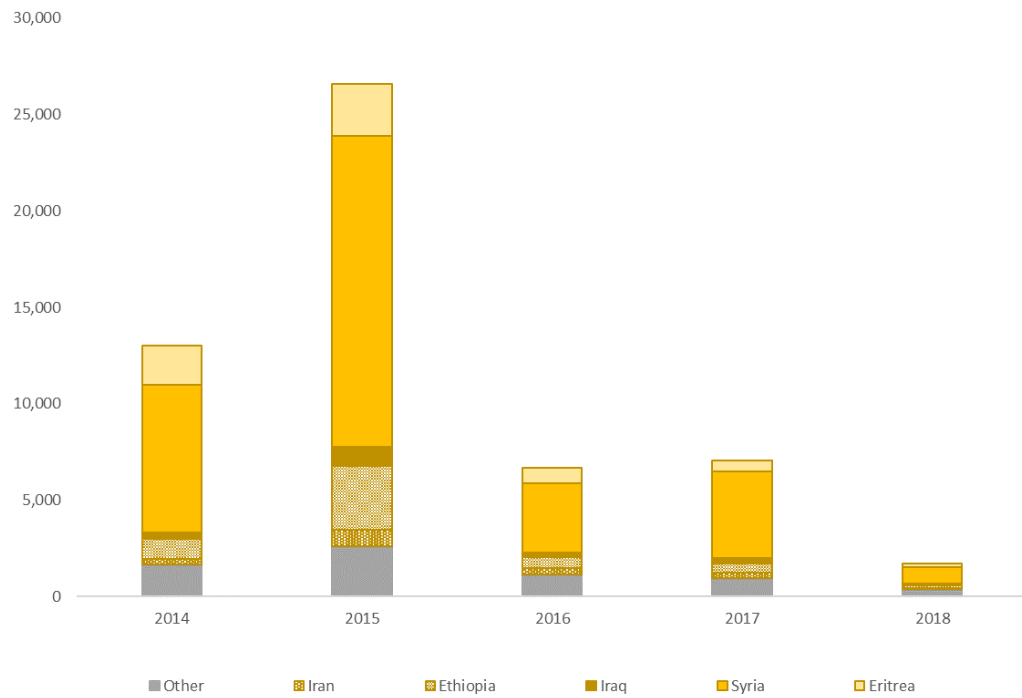
Notes: Column (2) presents who did not pass the civic integration exam nor naturalised. Column (3) presents those who passed the civic integration exam or were exempted. Hours worked in urban/rural, population density, co-national density, employed co-national density, and co-national stock are shown in 1000 scale.

naturalise either), those who did pass the exam, and those who fully naturalised. Differences in the characteristics of those who passed the exam and those who naturalised are small although the latter are more likely to be Syrian. Comparing those who naturalised or passed the exam with those who did neither, we see that the latter are more likely to be younger, female and to have education from the Netherlands (low or middle in particular), and less likely to be Syrian. These non-integrated individuals receive social benefits for a shorter duration and are less likely to be in employment. Again, there are few notable differences in terms of municipality characteristics.

Figure 1.3 provides the total number of refugees in our sample that arrived in each year and the countries of origin of the largest groups. Throughout the period, refugees from Syria are the largest group, although other countries account for a notable proportion in particular years too, e.g. Eritrea accounted for one-fifth of refugees in 2014.

Our study benefits from the quasi-random allocation of refugees to housing locations operated by the Dutch government's local units called COAs which are responsible for refugee housing. In the first instance, applicants are allocated to *AZCs* (refugee centres), located throughout the country, where they stay until they know whether they are granted asylum or have the right to stay in the Netherlands. The refugees are then moved into housing in municipalities. Unless there are exceptional circumstances, refugees have no say in the locations to which they are allocated and such allocations are proportionate to the population levels of

Figure 1.3: The number of asylum applications by arrival cohort



each municipality.[16]

Our data suggests that 87.2 per cent of refugees in the Netherlands remained within the location to which they were initially allocated after leaving the COA centre, for the duration of our sample period. Nevertheless to reduce concerns regarding sorting in our estimation sample we remove all ’movers’ and hence focus on those individuals who remained in their initially allocated locations.

[16] Exceptional circumstances include the need for medical treatment or the presence of immediate family members within a particular location. However, even if these criteria are met, in practice it is often impossible for the Dutch authorities to meet refugees’ needs or preferences due to long waiting lists for social housing. Furthermore, refusing social housing due to its location would result in a refugee remaining in the refugee centres (*AZCs*).

Table 1.2 shows which municipalities the sample refugees stayed in after holding an asylum residence permit (i.e. after they moved out of *POL* or *COL*). Approximately a quarter reside in relatively small municipalities, in which the largest COA centres, including *AZC*, are located. For instance, Westerwolde has a small population (25,000 as of 2019) yet contains the largest COA centre called Ter Apel with multiple functions of *COL, POL and AZC*. In Ter Apel, asylum seekers could go through the asylum application and stay and prepare for the integration exam after the asylum decision.[17] The locations of COA centres are also shown in Figure A.5 in Appendix.

Table 1.3 however suggests that large shares of refugee populations are located within the largest municipalities in our sample. These top five municipalities contain the largest Syrian communities in the Netherlands for instance (NL TIMES, 2016). However, when the integrated and the non-integrated groups are compared, there is only a small difference in current residential locations.[18] Due to the quasi-random allocation of refugees into *AZCs* and no substantial difference in current residential location by integration outcomes, the potential correlation of locational sorting with integration outcomes is unlikely to be a methodological concern in our analysis

[17] Not all of those who arrived in Ter Apel could stay after gaining the asylum residence permit.

[18] Table 1.1 also suggests no notable differences in population density between integrated and non-integrated refugees.

Table 1.2: Location of refugees by municipality

No	Municipality	Sample distribution	Population	COA facilities
1	Westerwolde	9.7%	25,199	2,000 shelters (*azc*, *col*, *pol*, other)
2	Cranendonck	5.0%	20,440	1,500 shelters (*azc*, *col*, *pol*)
3	Noordenveld	5.0%	31,290	- (*col*)
4	Noordoostpolder	3.3%	46,849	1,000 shelters (*azc*)
5	Dronten	3.0%	40,815	1,000 shelters (*azc*, other)

Note: Sample distribution provides the share of our sample who were allocated to each municipality while Population provides the population of each municipality as of January 2019 from StatLine. COA information as of 2021 COA (nd). A reception centre (i.e. *azc*) in Noordenveld closed in 2017 hence the lack of information for it.

Table 1.3: Municipal locations of refugees

No	Municipality	Sample distribution	Population distribution
1	Amsterdam	6.3%	5.0%
2	Rotterdam	4.0%	3.7%
3	's-Gravenhage	3.9%	3.1%
4	Utrecht	2.5%	2.0%
5	Groningen	1.6%	1.3%

Note: Population by municipality is as of January 2019 from StatLine.

1.5 Econometric Framework

We focus on legal integration in this study, assuming that gaining citizenship is a foundation for broader integration outcomes including participation in both labour markets and electoral processes.

1.5.1 Measurement of legal integration

Our analysis utilises two measures of legal integration. The first is whether or not individuals have passed the integration exam which is a necessary stepping stone to legal integration. The integration exam is a requirement for naturalisation and permanent residency and hence we do not know which of these routes successful examinees will choose. We consider gaining permanent residency to be a form of legal integration since permanent residents are entitled to almost the same rights as Dutch citizens including participating in most economic activities and voting in local elections. However, a crucial difference from naturalisation is that permanent residents are not allowed to vote in national elections (IND, 2019). Our second measure captures whether an individual became fully naturalised and recorded as a Dutch citizen in municipal registers. Due to the 5-year-residency requirement and the lengthy naturalisation application[19], those who obtained citizenship are confined to the 2014 cohort. Both of our integration measures are dummies, taking a value of 1 if the refugees are integrated by 2020 and zero otherwise.

[19] According to IND, the naturalisation application takes up to one year.

1.5.2 Model I - Role of individual characteristics on civic integration

This research extends the framework developed by Mossaad et al. (2018) to examine the effects of individual and location characteristics and the asylum application process itself on refugee integration in the Netherlands. A key advantage of our modelling framework is our ability to estimate the causal determinants of civic integration in the Netherlands due to the pseudo-random allocation of refugees across municipalities. More specifically we estimate the following equation using a probit model with standard errors clustered at the municipality level:

$$Pr(y_i = 1) = \alpha_i + \delta x'_i + \lambda econ'_i + \mu_m + \rho cont'_i + e_i \tag{1.1}$$

Where, y_i presents the integration outcomes of individual i by 2020, taking a value of 1 if the individual is integrated, 0 otherwise. x'_i denotes a vector of individual characteristics such as gender, country of origin, education enrolled in the Netherlands, and age at arrival.[20] We further include years since migration to the Netherlands (a measure of the number of years from arrival to integration or 2020 if no integration,), length of asylum application and arrival year dummies. $econ'$ is a vector of employment conditions including how long the individual received social benefits for and whether or not they are in employment. We further control for unobserved location fixed effects (FEs), μ_m. The municipality FEs are based on 355 mu-

[20] In this study, the top five largest refugee nationalities are coded as five individual dummy variables. All other nationalities are coded as ’others’ or ’stateless’.

nicipality categories capturing the municipality to which each individual was allocated and in which they therefore reside at the time of the exam. Finally, $cont$ indicates other control variables, which may affect integration outcomes, such as health conditions (i.e. mental health problems or chronic physical problems), and e is the error term. Appendix Table A.1 - A.3 provides a detailed description of all variables.

As mentioned in the previous section, we use two proxies for integration. When using the first measure, passing the integration exam, the full sample is used because everyone in the sample is eligible to take the exam (by construction). Hence our dependent variable is whether or not an individual passed the exam at any point between 2014-2020. When measuring integration in terms of naturalisation, the number of observations reduces to 9,482 as the refugees in this sample are from the 2014 cohort only, restricted by the minimum eligibility. The explanatory variable $arryear$ is removed from this specification.[21]

1.5.3 Model II - Role of assignment location on civic integration

After we isolate the effect of individual characteristics on civic integration, we focus on the impact of assignment location. Location is likely to be an important, yet unexplored determinant of the decision to integrate in a host country. We examine two specific channels

[21] In unreported estimations we also examine the likelihood of passing the exam using only the 2014 cohort to ensure this sample selection is not driving our results. Our findings in terms of sign and significance remain very similar to those estimated using the full sample.

relating to ’density’ and ’co-nationals’ to analyse how assignment locations associate with civic integration. First, following Duranton and Puga (2004), who show that learning, diffusion of knowledge and labour market interactions are higher in urban areas, we examine whether the likelihood of integration is higher when refugees are assigned to *denser* urban areas. Similarly, in line with De La Roca and Puga (2017), who point out the higher returns to urban market experience, we scrutinise whether refugees’ civic integration is influenced by the duration of employment in urbanised areas. In this specification we therefore include cumulative hours worked in urban and rural areas.[22] The cumulative experience gained in urban areas might be more valuable due to the richer opportunities provided by working at more productive firms and in higher quality skill-intensive jobs which could potentially speed up the likelihood to integrate. Also the effect of being already employed may impact integration. By taking these two mechanisms into account we try to show both a static treatment effect of location, and a dynamic effect that is working in cities increases the chances of better adaptation through matching and access to home country specific knowledge opportunities.

Second, following (Klaesson et al., 2021; Martén et al., 2019), we consider whether refugee integration is influenced by the presence of *co-nationals* in the assignment area. In line with Patacchini and Zenou (2012) who show that individuals, and particularly immigrants, are likely to associate with people who are similar, two types of mechanisms are expected to be in place. Primarily, a higher share or stock of co-nationals living nearby is a good

[22] The construction of cumulative hours of work is slightly different from the standard returns to big city arguments because refugees in our sample have never moved across areas of different density.

approximation of the possibility of social contacts among people from similar backgrounds, which may ease the adaptation of refugees in host countries and facilitate transmission and acquisition of information. More precisely though a higher share of *employed* co-nationals may provide an easier transition to labour market and/or finding jobs. We, therefore, scrutinise whether these mechanisms are detectable. However, while this presence of 'bonding social capital' within ethnic groups can prove beneficial, Putnam (2007) points out that the presence of co-nationals in an area may in fact be detrimental to social and civic integration if the group size is too large. In this situation the strong ties within the co-national groups can negatively impact on 'bridging social capital', thereby preventing individuals from building contacts and relationship with others. Indeed Battu et al. (2011) and Boeri et al. (2015) show a lower probability of being in employment due to the larger size of ethnic groups in the UK and Italy, respectively. However, as Beaman (2012) indicates, the empirical evidence for refugee groups is mixed. Therefore, we examine whether the presence of co-nationals and indeed employed co-nationals is an important factor for the civic integration of refugees and explore if there may be a threshold level of co-nationals by including quadratic terms in our specifications.

Finally, we examine whether the density and co-national channels differ with the level of education undertaken in the Netherlands by refugees. De La Roca and Puga (2017) find that more able workers enjoy greater learning advantages from bigger cities, suggesting that the effect of urbanisation on the likelihood to integrate may be greatest for the most educated refugees. Similarly, we explore whether the implications of co-national density for bonding and bridging social capital differ with refugees' levels of education.

To examine whether these locational characteristics influence integration outcomes we estimate the following equation:

$$Pr(y_i = 1) = \alpha_i + \delta x_i' + \lambda econ_i' + \gamma z_m' + \eta_c + \rho cont_i' + e_i \tag{1.2}$$

In Equation 1.2, z_m' is a vector containing socioeconomic variables at the municipality level[23] Municipal level variables include unemployment rate, population density per municipality, and co-national and employed co-national density per municipality. In these estimations, we replace municipal fixed effects with 40 labour market area fixed effects, η_c, known as COROP regions in the Netherlands.

Again, it should be underlined that endogeneity concerns relating to the higher density of people/co-national populations attracting refugees can arise when refugees are able to choose where they live. As mentioned, in our research design refugees are randomly allocated to refugee centres and housing across the country by the government. Furthermore, to remove the presence of non-systematic mobility towards more productive cities, our sample is re-

[23] As municipality boundaries have changed over time, these municipality variables are extracted based on the municipal jurisdictions in 2019 and 2020.

stricted to those who remained in the municipality where they were initially allocated until they passed the civic integration exam and/or naturalised.[24] To evidence the random distribution of the refugees in our sample, Appendix Table A.5 provides a summary of refugee characteristics for those allocated to urban settings and those to rural. As can be seen, the characteristics of refugees across the two assignment settings are very similar with no suggestion of a systematic assignment based on individual characteristics. This distribution suggests that residential location and integration outcomes are unlikely to be subject to reverse or two-way causality in this setting.

While sorting of the most able refugees to the cities should not therefore be present in our analysis, nor should there be particular concerns around measurement error of either integration or our key density variables, omitted variable bias can potentially impact our results. Although this should not be a significant concern since the Netherlands is a small country with similar locational endowments and natural resources, say *e.g.* weather or geology, we still try to address this point. We first include a wide selection of controls as well as municipality and time fixed effects in our specifications. Additionally, for any remaining effects of locational fundamentals that we were not able to control for, we instrument population density in baseline specifications. We follow Koster and Ozgen (2021) and use a historic measure of population density from 1902. Given more than 100 years time lag such an instrument should be uncorrelated with the current unobserved locational characteristics within

[24] We therefore remove the 12.8 per cent of the sample who moved from their initially allocated locations during the sample period. In unreported estimations we include these 'movers' in our estimations and find our key results remain unchanged.

our sample period, yet it exploits the strong autocorrelation of population density over time.

1.6 Results

1.6.1 Effect of individual characteristics

Before turning to our main interest, the effect of assignment location on civic integration, we first reconcile with the extant literature the effect of individual characteristics on civic integration.

Table 1.4 presents the baseline regression results based on Equation 1.1 where our *passed exam* measure of integration is the dependent variable. These estimations include all refugees in our sample who entered the Netherlands throughout our sample period. All specifications include arrival year and country of origin FE, family size and position in the household.

Column (1) of Table 1.4 provides the baseline specification with municipal fixed effects. Column (2) extends the specification by adding health and employment variables. Referring to the marginal effects reported in Appendix Table A.7, we see that the duration of the asylum application itself is associated with a reduced likelihood of passing the exam, with each additional month reducing the probability by 0.76 of a percentage point. For example, in our sample the average asylum application processing time is 4.5 months (with a 5.2 month standard deviation). This indicates that an applicant who had to wait a standard deviation longer to obtain their asylum permit would have a 4.0 percentage point lower likelihood of passing the exam.

Holding constant the duration of the asylum application, we find that an additional year spent in the Netherlands increases the probability of passing the civic integration exam by about 18 percentage points. This finding is in line with the refugee literature that indicates that the longer the time spent in the host country the higher the likelihood to integrate (Mossaad et al., 2018) and also with the broader literature on immigration.

As shown in column (2), the relationship between age and the probability of passing the integration exam is non-linear with the greatest likelihood of passing the exam being at the age of approximately 39. Moreover, females are less likely to pass the exam. Studying low/middle or high education in the Netherlands increases the likelihood of passing the exam with those enrolled in higher education being 12.9 percentage points more likely to do so, relative to those with no Dutch education. The impact of higher education is notably larger than that of low or middle education which increase the propensity to pass the exam by only 7.5 percentage points.

One of the most politicised topics about refugees is the relationship between economic self-sufficiency and integration. We show that being in employment at the time of taking the exam increases the likelihood of passing the civic integration exam by 8 percentage points. Similarly, the duration of the receipt of social benefits (as a share of time in the Netherlands) is positively associated with passing the exam. Finally, we find that refugees who suffer from mental health problems are 7.5 percentage points less likely to pass the exam while those with chronic physical problems are 8.2 percentage points less likely to do so.

Turning to our second indicator of integration, naturalisation, we find our results to be broadly similar to those in Table 1.4 and hence we report these results in Appendix Table A.9. The key differences are that the greatest likelihood of naturalising occurs at the age of 44.3 compared to 38.8 for passing the exam, and being female *increases* the likelihood of naturalisation but decreased the likelihood of passing the exam. The positive effect of being female on naturalisation is consistent with the findings of Mossaad et al. (2018).[25] Having low/middle education is associated with a reduced likelihood of naturalisation.[26] Finally the longer a refugee is in receipt of social benefits the less likely they are to naturalise which contrasts with the positive association between social benefits and passing the integration exam.

---

[25] For reasons of space we don't report the country of origin coefficients in Tables 1.4 and A.9 though the full results associated with Table 1.4 can be found in Table A.7. We find that refugees from Syria have an increased probability of passing the exam and naturalising relative to those from 'other' countries. In contrast, those from Ethiopia and Eritrea are less likely to pass the exam or naturalise. This may reflect the fact that neither Ethiopia or Eritrea allow dual citizenship and hence refugees from these countries may be reluctant to relinquish their Ethiopian or Eritrean citizenship in order to legally integrate.

[26] This finding is likely indicating that those with lower levels of education may take longer to navigate the complex bureaucratic processes associated with gaining citizenship and possibly are deterred by its financial cost. The current cost of the integration exam is 350 Euros, for which the individuals can use the social loan which is waived if the exam is passed within three years (DUO, 2022), compared to a minimum naturalisation cost of 925 Euros for a single person.

Table 1.4: Baseline regressions - individual characteristics (*Passed Exam*)

Dep Var: Passed Exam	(1)	(2)
Asylum application months	-0.0570***	-0.0372***
	(0.00158)	(0.00173)
Years since migration	0.844***	0.872***
	(0.0387)	(0.0415)
Arrival age	0.209***	0.139***
	(0.00560)	(0.00483)
Arrival age^2	-0.00275***	-0.00179***
	(7.20e-05)	(6.24e-05)
Female	-0.0553***	-0.0461***
	(0.0167)	(0.0174)
Low/Middle	0.129***	0.385***
	(0.0428)	(0.0348)
High	0.475***	0.714***
	(0.0577)	(0.0598)
Duration of social benefits		1.926***
		(0.0677)
Currently employed		0.405***
		(0.0259)
Mental health problems		-0.348***
		(0.0339)
Chronic physical problems		-0.373***
		(0.0635)
Municipality FE	Yes	Yes
Other controls	Yes	Yes
Observations	48,001	48,001
Pseudo R-sq	0.390	0.428

Note: Robust standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.1$. The reference groups for Education in the Netherlands, Arrival Year and Country of Origin are no education, 2014, and other countries, respectively. Other controls include family size, position in the household, stateless. Duration of social benefits is calculated based on the entire period of residence in the Netherlands (number of months to receive social benefits)/(number of total months since migration). The full tables, including average marginal effects, are shown in Appendix (Table A.7).

1.6.2 Effect of assignment location and co-national density

After setting the scene in terms of the role of individual characteristics for civic integration, we next turn to our *main* focus: the effect of assignment location - density and co-national channels - on integration outcomes. Our high-quality data and unique setting enables us to investigate which particular aspects of exposure to urban areas influence the integration outcomes of refugees. Table 1.5 reports the probit estimations for the civic integration exam of Equation 1.2 in which in addition to municipality level controls we include labour market area (COROP level) fixed effects. Table 1.6 repeats the same specification for naturalisation.

Panel A of Table 1.5, column (1) indicates a positive link between population density and the likelihood of passing the integration exam in line with our expectations, reflecting the fact that learning, diffusion of knowledge and labour market interactions are higher in urban areas (Duranton and Puga (2004)). The resources in urban areas available to refugees could indeed support their pace of learning to pass the exam. Table A.10 provides the average marginal effects associated with Table 1.5 and indicates that a doubling of population density from the mean level of 1,850 people per square kilometre would increase the likelihood of passing the exam by 4.1 percentage points (1.85*0.0222 = 0.041, where 1.85 is the sample mean level of population density in thousand/km^2).

However, excessive density may result in negative externalities. These may arise due to difficulties orienting in *large crowds* in big cities, often due to the impersonal ties one may have in dense urban areas which may lead to less attachment to the host country, but also in terms of the likely greater competition for the type of jobs immigrants typically do in the early years after arrival. These factors may result in less confidence in one's future prospects.[27]For example, Braun and Dwenger (2020) show in their study on expellees in post-war Germany that the more expellees that settled in a region, the lower the share of those who became economically integrated in the labour market, implying a reduced effort to integrate. Therefore, in column (2) we introduce the quadratic of the density variable. We show that the relationship between the civic integration exam and population density is inverted U-shaped with a turning point at approximately the 90^{th} percentile of population density within our sample. This finding indicates that while the relationship between population density and the likelihood of passing the exam is positive, the areas with very high levels of density seem to reveal negative externalities for integration.[28]

In the spirit of De La Roca and Puga (2017), albeit exploring a static return to urban experience, we further analyse whether a higher return to urban market experience increases the propensity to pass the exam. In other words, we focus on the value of the total duration of active participation in urban labour markets versus rural areas for civic integration.

[27] While our baseline specification does control for whether a refugee is currently in employment this work may be temporary or part-time in nature and hence wouldn't preclude a refugee from being concerned about the challenge of securing future employment.

[28] As previously discussed, the refugees in our sample had no influence over the locations to which they were dispersed upon leaving the reception centres so sorting of the most able to big cities should not be a major concern in our setting. Nevertheless, we instrument our population density and its quadratic term using historical measures of density from 1902 with results reported in Table A.13. The 2SLS results are consistent with those already reported. Therefore, given the limited number of observations in the following sub-sample estimations and efficiency concerns, we proceed with presenting probit results throughout. The 2SLS results are available upon request.

In order to explore the effect of employment duration on the likelihood of passing the exam we first need to address the fact that those who don't pass the exam will, by the end of our sample period, have been in the Netherlands, and hence had the opportunity to work, longer than a refugee who passed the exam at some point within the sample period. To do so we censor our data to create a sub-sample of refugees who have all been in the Netherlands between 3 years and 6 months and 3 years and 8 months which is one month either side of the mean time taken to pass the exam in our sample. By restricting our sample to those who have all been in the Netherlands for an identical length of time, we can then compare the effect of employment duration on their likelihood to pass the exam during this specific two month period.[29]

Column (3) of Table 1.5 focuses on this sub-sample and includes the duration of total employment in urban and rural areas separately. It should be emphasised that none of the refugees in our sample changed their initial assignment location, thus in our setting there are no refugees who worked both in urban *and* rural areas. Our results indicate that the duration of employment does have a positive association with the likelihood of passing the civic integration exam for both urban and rural employment. The marginal effects in Table A.10 indicate that an increase in hours worked equivalent to a year of full-time employment (approximately 1,800 or 1.8 when scaled in thousands of hours) would increase the likelihood of passing the integration exam by 2.4 percentage points (0.0133*1.8) for urban workers and

[29] The refugees in both control and treatment groups are very similar in terms of their observational characteristics examined in Table 1.4.

2.3 percentage points for those in rural employment (0.0129*1.8).

In Panel B, we analyse whether the residential proximity of co-nationals increases the likelihood of refugees to pass the exam by providing support networks and the diffusion of knowledge which could lead them to settle and integrate sooner in the host country. Column (1) shows that there is a positive link between the density of co-nationals in a municipality and passing the integration exam. From the marginal effects in Table A.10 we calculate that a doubling of co-national density from the mean level of 40 co-nationals per square kilometre would increase the likelihood of passing the exam by 1.0 percentage point. Column (2) indicates that this effect is non-linear with a turning point approximately equal to the 95^{th} percentile of co-national density. This finding is consistent with the effects of bonding social capital beginning to adversely impact bridging social capital at high levels of co-national density thereby reducing the probability to integrate in the host country (Putnam, 2007; Boeri et al., 2015). Column (3) includes population density in the same specification to isolate the potential correlation between population density and the density of co-nationals since dense urban areas also tend to contain large numbers of foreign populations. Our co-national result survives this check, and hence indicates that co-national density has a positive effect on integration outcomes, holding urban density constant.

Table 1.6 repeats Table 1.5 for our second measure of legal integration, that is naturalisation. These estimations are based only on the 2014 cohort which means we lose more

Table 1.5: Baseline regressions - municipal level (*Passed Exam*)

Dep Var: Passed Exam	(1)	(2)	(3)
Panel A: Density Channel			
Population density	0.109***	0.235***	
	(0.0296)	(0.0604)	
Population density^2		-0.0212***	
		(0.00666)	
Hours worked in urban			0.288***
			(0.0827)
Hours worked in rural			0.278***
			(0.0969)
Observations	48,001	48,001	7,014
Psuedo R-sq	0.430	0.431	0.832
Panel B: Co-national Channel			
Co-national density	1.264***	3.675***	2.155***
	(0.344)	(0.608)	(0.511)
Co-national density^2		-7.178***	-3.751***
		(1.372)	(1.515)
Population density			0.198***
			(0.0623)
Population density^2			-0.0199***
			(0.00720)
Observations	48,001	48,001	48,001
Psuedo R-sq	0.429	0.430	0.432

Note: Robust standard errors in parentheses, $***p < 0.01, **\ p < 0.05, *p < 0.1$. Estimations contain all variables in Table 1.4, as well as unemployment rates and COROP fixed effects. Average Marginal Effects are presented in Appendix Table A.10.

than three quarters of our observations. In Panel A, population density exhibits a positive effect on naturalisation though this effect is less precisely estimated than the effect on the civic integration exam. We again show a non-linear effect of population density with respect to the likelihood of naturalisation. The effect becomes negative beyond the 95th percentile of population density, *e.g. Amsterdam*, which echoes the result found for passing the civic integration exam. Referring to the average marginal effects in Table A.11, a doubling of the mean level of population density would result in a 2.7 percentage point increase in the likelihood of naturalising. The effect of population density on naturalisation is 35% less than its effect on the civic integration exam, suggesting the role of the assignment location is less prominent for naturalisation.

In column (3) we estimate the impact on naturalisation of the duration of employment in urban versus rural areas. Since our naturalisation sample focuses only on our 2014 cohort and examines whether or not they have naturalised by 2020 we don't have the same challenge that our passed exam analysis faced whereby those who did not pass the exam were likely to have greater employment duration than those who did. Column (3) in panel A shows that accumulated experience in urban areas has a positive effect on naturalisation that is larger and more precisely estimated than for those employed in rural areas. The marginal effects in Table A.11 indicate that an increase in hours worked equivalent to a year of full-time employment would increase the likelihood of naturalising by 1.2 percentage points (0.00694*1.8) for refugees with accumulated labour market experience in urban areas.

Panel B Table 1.6 indicates a non-linear relationship between co-national density and naturalisation which is consistent with the results for the civic integration exam. However the effect is less precisely estimated for naturalisation with the standard errors associated with the average marginal effects in Table A.11 being larger. We find that doubling co-national density at the mean will increase the likelihood of naturalising by 1.4 percentage points.

Table 1.6: Baseline regressions - municipal level (*Naturalisation*)

Dep Var: Naturalised	(1)	(2)	(3)
Panel A: Density Channel			
Population density	0.0103	0.147***	
	(0.0227)	(0.0496)	
Population density^2		-0.0224****	
		(0.00672)	
Hours worked in urban			0.0305**
			(0.0143)
Hours worked in rural			0.0245
			(0.0166)
Observations	9,482	9,482	9,482
Psuedo R-sq	0.171	0.172	0.171
Panel B: Co-national Channel			
Co-national density	0.0734	2.002**	1.547
	(0.337)	(0.906)	(0.964)
Co-national density^2		-5.569**	-4.022
		(2.422)	(2.510)
Population density			0.121**
			(0.0523)
Population density^2			-0.0193***
			(0.00708)
Observations	9,482	9,482	9,482
Pseudo R-sq	0.171	0.171	0.172

Note: Robust standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.1$. Other controls contain all variables in Table A.9, as well as unemployment rates and COROP fixed effects are included. Average Marginal Effects are presented in Appendix Table A.11).

Our data do not allow us to observe pre-arrival personal characteristics. However, similar to Eckert et al. (2022) we do observe post-arrival educational attainment. Obtaining a host country education degree may not only be indicative of a refugee's ability but may also point to their intention to invest in the host country long term. Therefore, in Table 1.7 we explore whether the effects of urbanisation and co-national density differ with the level of education of refugees. Columns (1) and (2) show that the effects of population and co-national density on the likelihood of passing the exam are larger for those with high levels of education than those with no or low education. Referring to Table A.10, we find that the effect of doubling mean population density results in an increase in the likelihood of passing the exam by 4.7 percentage points for the no/low educated and 5.8 percentage points for the highly educated. Comparable figures for the effect of co-national density are 1.5 and 2.3 percentage points, respectively.[30]

Turning to naturalisation, columns (3) and (4) indicate that the effects of co-national density are largest for the highly educated which is in line with our findings for the integration exam. However, a degree of caution is warranted for these estimations given the substantial reduction in the sample size. For naturalisation, we are able to explore how the effects of urban and rural employment duration differ by education level.[31] To do so, columns (5) and (6) of Table 1.7 show the effect of hours worked in urban and rural areas for sub-samples containing only those with no or low education (column (5)), and only those with high education (column (6)). Consistent with the findings of Eckert et al. (2022), we show that hours worked in urban areas have a stronger effect on the likelihood of obtaining citizenship for the highly educated. The effects of employment in rural areas is estimated positive for no/low-educated and even negative for the highly-educated. However, these coefficients are imprecisely estimated.

[30] e.g. 1.5 percentage point increase is calculated as [0.426-2(0.734*0.04)]*0.04 = 0.0146, where 0.04 is the mean value of co-national density.

[31] We are not able to estimate comparable results for passing the exam as the sample size for the highly educated becomes too small with almost no variation when the sample is censored to contain only those who were in the Netherlands between one month before and after the mean exam passing time in our sample.

In sum, the results of Table 1.7 provide some evidence to suggest that the effects of urban employment duration, population density and co-national density on the likelihood of integration are greatest for those with high levels of Dutch education. This may reflect the fact that more able refugees are better able to take advantage of the opportunities provided by greater social interaction, knowledge transmission and more productive employment.

1.6.3 Further Exploration and Robustness checks

To further explore our results and to assess if they are sensitive to a number of definition, measurement and sample selection issues, we undertake several robustness exercises.

First, variables such as population density, co-national density and total hours worked are often transformed in some way to address the fact that they tend to be skewed to the right. Since natural logs are not defined at zero, a recent and commonly used approach, including in migration studies (e.g.Bahar and Rapoport (2018)) is to transform variables using the inverse hyperbolic sine (IHS) method.[32] Since this approach is sensitive to the scale of variables (Norton, 2022) in our main results we prefer to use non-transformed measures of these variables. Nevertheless, here we report results in which population density, co-national density and hours worked are transformed using the IHS method. Table 1.8 provides the results for passed exam and naturalisation and indicates that the sign and significance of population den-

[32] The IHS transformation of a variable x is equal to $\log(x + (x^2 + 1)^{1/2})$.

sity, co-national density and hours worked remain almost identical to those estimated using non-logged versions of these variables. Again, turning points for population and co-national densities are estimated at the upper end of the sample range.

Second, we consider alternative ways to capture how the presence of co-nationals affects the likelihood of integration. Until now we have suggested that the bonding social capital that may influence integration is best captured using a measure of the *density* of co-national populations. However, it could also be argued that the absolute numbers of such populations are also important because what matters is more the exposure to co-national populations in an area which need not be correlated with the size of that area. Panel A of Table 1.9 therefore provides our results for the civic integration exam (columns 1-3) and naturalisation (columns 4-6), where we replace co-national density with a measure of co-national stock (total number of co-nationals in a municipality). For the civic exam, again the sign and significance remains unchanged with a turning point above the 90^{th} percentile of co-national stock. For naturalisation the effect of population density remains in line with the baseline estimation, while the co-national density results become imprecise though positive.

Next, following Klaesson et al. (2021) we distinguish between co-national density and *employed* co-national density on the basis that it could in fact be employed co-nationals rather than co-nationals *per se* who are most likely to assist with networking, gaining employment and perhaps ultimately legal integration. Panel B of Table 1.9 therefore explores how employed co-national density affects the likelihood of passing the civic integration exam and naturalising. Both forms of integration can be seen to have an inverted U-shaped association with employed co-national density, again with turning points at the very upper end of the distribution. Estimated marginal effects are comparable to those for total co-national density. For instance, a doubling of employed co-national density from the mean would increase the likelihood of naturalising by 1.6 percentage points, while a doubling of *total* co-national density would increase the likelihood of naturalising by 1.4 percentage points.[33].

Finally, given that our sample contains not only those who passed the civic exam but also those who were deemed to be exempt from it, we here utilise an alternative sample in which we exclude those who gained exemption from the exam in case they are not representative of those who actually sat the exam. Figure A.6 graphically illustrates our key density and co-nationality results for our base sample and the alternative sub-sample. It is clear that there is very little variation in sign, significance or magnitude of the estimated coefficients across the two samples.

[33] Comparing the model in column (2) of panel B in Table 1.6 with that from column (5) of panel B in Table 1.9

Table 1.7: Baseline regressions by education groups - municipal level

	Passed Exam		Naturalisation			
	No & Low	High	No & Low	High	No & Low	High
	(1)	(2)	(3)	(4)	(5)	(6)
Population density	0.195***	0.480*			0.125**	-0.490
	(0.0615)	(0.251)			(0.0538)	(0.784)
Population density^2	-0.0192***	-0.0769**			-0.0203***	0.0744
	(0.00711)	(0.0347)			(0.00727)	(0.106)
Co-national density	2.084***	4.158*			1.381	10.43
	(0.519)	(2.341)			(1.003)	(19.17)
Co-national density^2	-3.589**	-6.577			-3.839	29.62
	(1.531)	(5.671)			(2.647)	(85.56)
Hours worked in urban			0.0267*	0.291		
			(0.0149)	(0.198)		
Hours worked in rural			0.0240	-0.131		
			(0.0171)	(0.235)		
Observations	47,019	957	9,269	134	9,269	134
Pseudo R-sq	0.432	0.481	0.170	0.415	0.171	0.488

Note: Robust standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.1$. Other controls contain all variables in Table 1.4, as well as unemployment rates and COROP fixed effects are included.

Table 1.8: Robustness check: Inverse Hyperbolic Sine Transformation

	Passed exam			Naturalisation		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Density Channel						
Population density	0.292***	0.342***		0.0955*	0.332***	
	(0.0790)	(0.123)		(0.0567)	(0.124)	
Population density^2		-0.0202			-0.0958**	
		(0.0329)			(0.0432)	
Hours worked in urban			0.496***			0.0687**
			(0.144)			(0.0298)
Hours worked in urban			0.437***			0.0560
			(0.154)			(0.0354)
Observations	48,001	48,001	7,014	9,482	9,482	9,482
Pseud R-sq	0.431	0.431	0.832	0.171	0.172	0.171
Panel B: Co-national Channel						
Co-national density	1.299***	3.744***	2.336***	0.0835	2.059**	1.806*
	(0.347)	(0.627)	(0.500)	(0.345)	(0.926)	(0.960)
Co-national density^2		-7.487***	-4.295***		-5.868**	-4.922**
		(1.455)	(1.516)		(2.547)	(2.563)
Population density			0.341***			0.300**
			(0.130)			(0.125)
Population density^2			-0.0475			-0.0936**
			(0.0350)			(0.0437)
Observations	48,001	48,001	48,001	9,482	9,482	9,482
Pseud R-sq	0.429	0.430	0.432	0.171	0.171	0.172

Note: Robust standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.1$. All reported variables are transformed using the Inverse Hyperbolic Sine transformation. Other controls, unemployment rates and COROP fixed effects are included.

Table 1.9: Robustness check: Alternative measures of co-national density

	Passed exam			Naturalisation		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Co-national stock						
Co-national stock	0.00878***	0.0239***	0.0161***	0.00343**	0.000848	0.000585
	(0.00140)	(0.00503)	(0.00435)	(0.00169)	(0.00747)	(0.00681)
Co-national stock^2		-0.000290***	-0.000153**		4.94e-05	7.19e-05
		(8.05e-05)	(7.34e-05)		(0.000132)	(0.000118)
Population density			0.238***			0.162***
			(0.0602)			(0.00499)
Population density^2			-0.0254***			-0.00250***
			(0.00705)			(0.00681)
Observations	48,001	48,001	48,001	9,482	9,482	9,482
Pseud R-sq	0.430	0.430	0.433	0.171	0.171	0.172
Panel B: Employed co-national density						
Employed co-national	4.287***	12.51***	7.778***	0.250	7.617***	6.108*
	(1.065)	(2.028)	(1.814)	(1.201)	(2.882)	(3.118)
Employed co-national^2		-82.82***	-52.16***		-75.41***	-58.41**
		(15.44)	(18.81)		(25.01)	(27.77)
Population density			0.190***			0.117**
			(0.0608)			(0.0512)
Population density^2			-0.0185***			-0.0184***
			(0.00702)			(0.00697)
Observations	48,001	48,001	48,001	9,482	9,482	9,482
Pseud R-sq	0.429	0.430	0.432	0.171	0.172	0.172

1.7 Discussion and Conclusions

When refugees do not achieve full legal integration in host countries they become socially and economically isolated and legally invisible to politicians with no right or ability to influence policy. Yet despite the complexity and financial cost of refugee settlement programmes and the economic and social impacts that would arise from their failure, policy makers currently have little guidance as to what factors may hinder or facilitate refugee integration. Existing evidence tends to be US based, focuses exclusively on labour market integration and/or examines earlier waves of refugees into Europe which were different in terms of both scale and ethnic composition. With this background in mind we examine rich, individual-level data for refugees who arrived in the Netherlands between 2014 and 2020 to identify the factors that affected their likelihood of achieving legal integration along two dimensions. These are passing the civic integration exam that allows one to obtain a permanent residency in the Netherlands; and full naturalisation, hence obtaining Dutch citizenship. Benefiting from the quasi-random allocation of refugees within the Netherlands, we are able to identify the role played by the asylum process, and both individual-level and assignment location characteristics. The latter allows us to specifically focus on how urban density and co-national channels may have influenced refugee integration.

We find that the individual characteristics of refugees affect civic integration outcomes significantly. However, holding these characteristics constant, we also show that the level of urbanisation in the assignment location is an important determinant of civic integration.

Refugees located in densely populated areas are more likely to take and pass the civic integration exam and ultimately to naturalise than those in less dense areas. However, the marginal effect of urbanisation on integration declines as urbanisation increases, and becomes negative at high levels of urban density. This suggests the presence of negative externalities associated with high levels of density which reduce the likelihood of integration. We also find evidence to indicate that the *duration* of employment in urban areas increases the likelihood of integrating, while a greater length of time working in rural areas has a smaller effect. We also find that the positive relationship between integration and urbanisation, and the duration of employment in urban areas, is greatest for those with higher education. These findings are in line with those of Eckert et al. (2022) and (De La Roca and Puga, 2017) who argue that learning, the diffusion of knowledge and labour market interactions are all greatest in urban areas. (De La Roca and Puga, 2017) additionally find that the benefits of big city experience is greatest for those with higher initial ability, again consistent with our findings.

We further show that co-national density is an additional mechanism that affects integration, beyond density. This finding supports earlier work which indicates that the presence of co-nationals increases social contacts, eases adaptation to the host country and enhances the transmission and acquisition of knowledge, all factors that may be expected to increase the likelihood of an individual to integrate. However, we do find evidence of an optimal level of co-national density beyond which further increases in density make integration less likely. This finding is consistent with the argument that beyond this point co-national density begins to adversely affect bridging social capital, with co-national groups perhaps becoming more insular, more self-sufficient, and less likely to integrate (Putnam, 2007).

In terms of individual characteristics, we find that over the period 2014-20 those passing the civic integration exam were most likely to be male, aged around 39, and from Syria. Those who became naturalised were more likely to be female, aged 44 and also from Syria. Our results indicate that a number of factors reduce the likelihood of refugees obtaining either form of legal integration. These include personal characteristics such as suffering from chronic physical or mental health conditions as well as aspects of the asylum process such as its duration. The likelihood of gaining naturalisation is further reduced by the duration of time an individual spends in receipt of social benefits. In contrast, we find that the likelihood of both forms of legal integration is increased by an individual being in employment and being educated within the Netherlands.

Our results should prove useful to those involved in refugee settlement policy and provide several direct policy implications. First, locating refugees in low population density areas is likely to result in lower rates of integration. If this cannot be avoided then additional support and opportunities may be required to ensure that refugees are not missing out on opportunities for knowledge diffusion, networking, and employment. Relatedly, locating refugees within large co-national communities may also hinder integration without compensating efforts to improve bridging social capital. Second, providing refugees with both employment and educational opportunities will have a clear benefit to their prospects for legal integration. While gaining human capital and experience of employment are important components of economic integration, their benefits are likely to go beyond this. Attending school and participating in the workforce are likely to also improve language skills and knowledge of host country culture, practices and institutions and thereby significantly boost social integration. Third, a perhaps neglected deterrent to integration is the physical and mental health of refugees. Increasing awareness of, and access to, appropriate health services for refugees is therefore likely to yield wider, longer-term societal benefits beyond the health of the individual refugees. Finally, our results consistently show that the duration of the asylum application process reduces the likelihood of refugees passing the integration exam and gaining full naturalisation. As with healthcare, additional resourcing of the asylum process is therefore likely to yield wider benefits to host countries.

Since any policies that can reduce the obstacles to refugee integration are likely to be of significant long-term benefit to host countries, these policy implications should provide some ammunition to policymakers who wish to overcome populist opposition to further investment into the asylum process. Only with full legal integration can refugees fully contribute to host country economies, participate within their societies and engage with and become *visible* to and be heard by politicians and policymakers.

Chapter 2

Local Labour Market Training and Civic Integration Exams: Examining the Employment Outcomes of Refugees

Abstract

We explore whether a civic integration component dedicated to labour market training (the ONA) boosts refugees' economic integration. Using linked employer-employee data from 2014 to 2021 for the Netherlands, we compare the labour market trajectories of refugees required to take the ONA component with those who were not. We find employment probability to be 6.4-11.0 percentage points higher for those who took the ONA but find no significant differences in hours worked or hourly wages. We also explore the mechanisms and firm effects through which the ONA operates. Finally, the ONA doesn't improve the persistent labour market disadvantages faced by females and Syrians.

Keywords: Refugees, labour market performance, Integration exam, Employment, Netherlands.

JEL classification codes: J08, J15

2.1 Introduction

The Arab Spring movement and the resultant instability within Syria and other Arab nations caused a significant inflow of asylum seekers into European countries. Indeed, since 2014 the number of refugees seeking asylum in Europe has increased steadily, reaching 6 million by 2019 (UNHCR, 2021). With the ongoing war in Ukraine leading to further inflows of refugees to European countries, a key challenge faced by policymakers is to better understand how host countries can integrate refugees into their societies and economies. A failure to integrate refugees is likely to result in significant costs to the welfare system, widened inequality and potentially higher crime rates (Enchautegui and Giannarelli, 2015; Adamson, 2020; Foged et al., 2022). In order to facilitate the integration of refugees, several European countries have implemented integration schemes that combine support programmes such as local language courses with legal requirements to pass civic integration exams and/or to declare respect for local social norms and values. Despite these integration programmes being costly, their effectiveness has received relatively little attention.[1] While it's widely agreed that a key aspect of successful integration is labour market participation, policymakers currently have little guidance as to which components of civic integration training are beneficial to the economic integration of refugees.

The aim of this paper is to fill this gap in the literature by assessing the impact of a unique local labour market training programme on the economic integration of refugees. In con-

[1] Goodman and Wright (2015) explore the broader socio-economic and political effects of civic integration requirements on immigrants, as opposed to refugees, finding them to have little tangible impact.

trast to civic integration programmes applied elsewhere, a policy change in the Netherlands in 2015 introduced dedicated labour market training that is practical in nature and requires refugees to actively demonstrate the knowledge acquired in order to pass this component. Our analysis, which follows the universe of refugees and their employers over time, also allows us to explore refugee and firm heterogeneity and the *mechanisms* through which this policy affects labour market outcomes. The rate of asylum applications in the Netherlands, at approximately three per 1,000 population, is broadly representative of the wider European Union. Asylum applicants are required to pass a complex set of exams including, since 2015, a component specifically designed to improve their knowledge and understanding of the Dutch labour market and to provide practical, hands-on training to identify, apply for, and secure suitable jobs. This component incorporates CV writing, participating in mock job interviews and other labour market training. Facing large numbers of asylum applications - over 45,000 in 2015 alone - and with a well-developed and unique civic integration programme, the Netherlands provides a good setting to test the effectiveness of such an integration requirement.

Our analysis contributes to a growing strand of literature that focuses on the integration of refugees in host countries and documents the specific integration investments that aim to improve their labour market outcomes. Refugees tend to face persistent and long-term disadvantages in comparison to other immigrants in terms of employment rate, income and occupational quality (Bevelander, 2020). For example, Fasani et al. (2022) estimate that the unemployment rate of refugees may be 22% higher than that of other migrants with similar characteristics and fails to converge even 10 years after migration. Similarly, Müller et al. (2023) identify gaps in employment rates between refugees and native workers which are not explained by human capital or other characteristics and that persist for two decades. The labour market disadvantages experienced by refugees have a number of causes including the presence of physical and/or mental health conditions resulting from traumatic experiences in their home countries, lack of preparation to move into the host country, forcibly disrupted education or working tenures, and delayed participation in host country labour markets due to legal restrictions and discrimination (De Vroome and Van Tubergen, 2010; Kone et al., 2019; Brell et al., 2020). Moreover, some studies highlight a heterogeneity in labour market outcomes among refugee groups. For example, female refugees experience lower rates of labour market participation and refugees from Syria and Somalia face a higher risk of long-term unemployment (Helgesson et al., 2019; Bevelander, 2020).

The integration outcomes of refugees depend heavily on government policies and programmes, such as asylum procedures, allocation to public residential facilities, restricted socio-economic activities such as working or voting, as well as support including social benefits and language training. Our paper contributes to a growing literature on active labour market policies targeted at refugees. For example, Arendt (2022) explores the impact of on-the-job training implemented in Denmark and shows a short-term positive effect on labour market participation for male refugees. Battisti et al. (2019) conduct a field experiment on approximately 400 refugees in Germany in which some are offered job-matching support. They find positive effects on employment a year after the treatment. Finally Dahlberg et al. (2022) evaluate a randomised control trial targeted at 140 low educated refugees in the city of Gothenburg in Sweden and show that early and intensive training has significant positive effects on unemployment after 12 months.[2] Key limitations of the studies in this area to date are the relatively short time horizon over which the impact of policies are assessed, an inability to link refugees with firm-level data, and the small numbers of refugees included in experimental studies. These are all limitations that our paper addresses.

A number of studies explore other ways in which government policy may influence the economic integration of refugees. Fasani et al. (2021), for example, show that banning working during the period of asylum claims results in lower labour market participation for up to 10 years post-arrival. In addition, Hainmueller et al. (2016) find that lengthy asylum applications have an adverse effect on mental health leading to negative implications for labour market outcomes. Aslund et al. (2022) also provide evidence of negative impacts of longer waiting on labour market experiences in Sweden. It has also been suggested that the policy of geographic dispersal of refugees harms their economic performance by limiting their access to (employed) immigrant and/or co-ethnic networks (Edin et al., 2004; Godøy, 2017; Martén et al., 2019). In contrast, evidence suggests that the appropriate provision of rights and opportunities can help refugees to catch up in the labour market. Hainmueller et al. (2019)

[2] Focusing on immigrants rather than refugees, Sarvimäki and Hämäläinen (2016) find that an active labour market policy increased the earnings of 1990s immigrants by 47% over a 10 year period.

provide causal evidence that gaining citizenship leads to additional annual earnings of 5,000 US Dollars for the subsequent 15 years, while other studies, such as Lochmann et al. (2019), indicate positive effects of language training. For instance, Battisti et al. (2019) demonstrate the importance of job search assistance for refugees in Germany, while Foged et al. (2022) show that access to local language training improves employment and earning outcomes and upgrades refugees' occupations to complex jobs with stronger communication requirements. Language training is thought to offset some of the labour market disadvantages caused by a mismatch between the local spoken language and refugees' language skills (Auer, 2018). Foged et al. (2022) also show that lowering welfare benefits does not provide an incentive to work.

Despite the growing academic interest in examining refugee integration, the direct role played by civic integration exams on labour market outcomes has yet to be fully assessed. Although language and integration programmes are common, they are frequently passive in nature, rather than requiring practical demonstration of the learning. In 2015, in addition to their existing civic integration programme, the Netherlands introduced a unique exam component called Orientation about the Dutch Labour Market (ONA) to assess immigrants' knowledge of the Dutch labour market and their readiness to find and apply for suitable jobs. This additional component is a unique programme that goes a step beyond standard integration exams and tries to improve *employment-related* linguistic skills combined with knowledge and training relating to the particulars of the Dutch labour market. All refugees granted asylum since January 2015 are required to pass the standard civic integration exam *and* the ONA component as part of the requirements for obtaining permanent residency in the Netherlands. In contrast to Arendt et al. (2022)'s study of Denmark, this labour market training is *in addition* to linguistic training and requires the active use of the Dutch language through preparing a CV and having an in-person mock job interview in Dutch for a real vacancy in order to pass the integration exam.

An important contribution of this paper, therefore, is to utilise this unique policy setting to examine and compare the effects of labour market training (i.e. the ONA) as a part of civic integration tests on refugees' labour market outcomes. Our identifying variation comes from comparing the trajectories in labour market performance of two groups of refugees: first, those who are obliged to complete the mandatory ONA component as part of the post-2015 civic integration exams; and second, those who take the pre-2015 civic integration exam *without* this component. We use high-quality linked employer-employee administrative data for the Netherlands for the period 2014 to 2021 and apply Propensity Score Matching (PSM) to mitigate potential biases caused by compositional differences between the two groups. Importantly, our analysis examines the universe of all refugees in the Netherlands over an 8 year period. We also follow refugees over time allowing us to control for their unobserved abilities and pre-arrival human capital endowments. By linking refugees to the firms that they work for we also explore the role of firms on refugee employment outcomes. In this manner, we believe we make a clear contribution to the literature examining the impact of active labour market policies on the economic integration of refugees.

Our results indicate that those who completed the ONA component out-perform those who did not in terms of labour market participation after a period of approximately four years from the initial asylum application.[3] The divergence between the two groups increases from that point and reaches 6.4 percentage point difference in the probability of employment after eight years. This difference reaches 11 percentage points for male refugees across the two cohorts. However, the positive impact of the ONA on employment does not translate into a statistically significant difference between these two groups in terms of higher hours worked and hourly wages and is robust to controlling for unobserved firm effects. Hence, our baseline findings indicate that the ONA component does not effectively bridge the human capital gap that hinders refugees from accessing higher-paying employment opportunities. Furthermore, our analyses of firm heterogeneity reveal no significant statistical distinction between refugees with and without ONA training in terms of their employment in higher value-added and lower labour share firms. These results further support the notion that the ONA programme fails to address the underlying factors that contribute to refugees' limited access to more lucrative job prospects.

After exploring potential mechanisms -learning, signalling, networking and geographic mobility- behind this employment effect of the ONA, we conclude that the ONA effect oper-

[3] In the Netherlands, refugees can work once they are granted temporary residence permit and the average time to pass the civic integration exam is around three years.

ates through longer employment duration and fewer job changes, which we label as a learning effect. This means the ONA training leads to higher job stability through improved matching, but does not lead to employment in higher quality jobs. Our results indicate that those who passed the ONA component changed jobs less frequently (on average they held 1.0 fewer job over our sample period) and were significantly less likely to be mobile geographically, while their average duration of employment was 3 months longer, compared to those who did not pass the ONA component.[4] We also find that employers do not view passing the ONA as a signal of superior skills.

Controlling for human capital and country of origin, we show that the impact of the ONA is not symmetric across different groups of refugees; female refugees and refugees from Syria remain at a disadvantage in the labour market even after the completion of the ONA. The positive effect on employment probability is mostly driven by male refugees and refugees from countries other than Syria. Since Syrian refugees form the largest cohort in the Netherlands in our period of analysis this implies that large numbers of individuals are not benefiting from the ONA, suggesting the need for a targeted policy response.

The remainder of the paper is structured as follows: Section 2 provides an overview of integration policies in the Netherlands; Section 3 presents the data and details the construction of our sample; Section 4 outlines our empirical approach; Section 5 discusses our findings;

[4] Sample means are 2.5 jobs and 8.9 months of employment duration.

Section 6 provides robustness check results, and Section 7 concludes.

2.2 Civic Integration in the Netherlands

In common with other European countries such as the United Kingdom, France and Germany, the Netherlands requires asylum migrants (i.e. refugees) to fulfil the requirements of a civic integration programme in order to be able to reside in the country. These requirements aim to cultivate basic language proficiency and knowledge of Dutch society and to ensure that refugees are not economically or socially marginalised (Enchautegui and Giannarelli, 2015; Department of Dutch, UCL, nd). After obtaining an Asylum Residence Permit, refugees in the Netherlands are housed across the country on a quasi-random basis and are required to pass the civic integration exam. In this section, we overview the evolution of Dutch integration policies over time and outline the components of the civic integration programme that current refugees are required to complete.

2.2.1 Changes in Integration Requirements

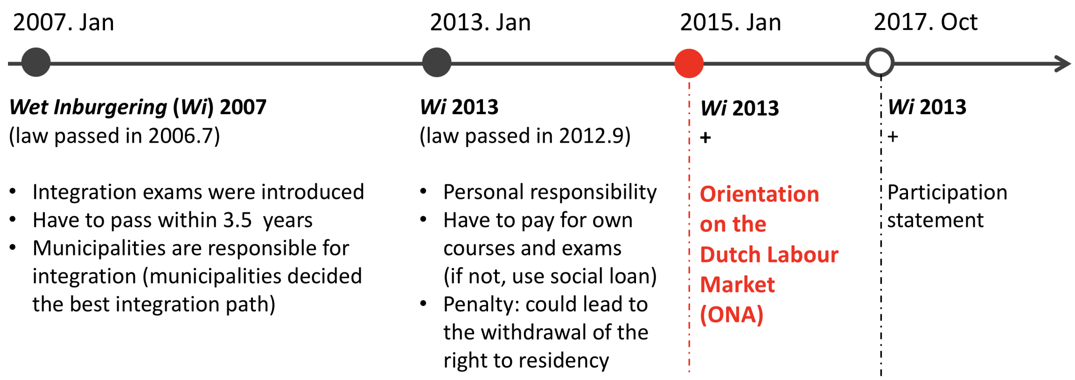
Before 2006 asylum applicants were required to take part in a relatively 'light touch' integration programme which contained no formal testing or assessment. However, the Civic Integration Act of 2007 (*Wet inburgering* 2007: *Wi*2007) required asylum migrants arriving from January 2007 to pass a civic integration exam within a period of three and a half years from obtaining a residence permit. The integration process was the responsibility of the municipality where immigrants resided, with the municipalities expected to support immigrants in this process and to recommend the best integration path through the programme (Algemene Rekenkamer, 2017).[5]

In 2013 an amended act (*Wi*2013) came into effect which aimed to reduce the financial burden on local governments, provide a variety of course options, and shorten the integration process (Henskens and Dijsselbloem, 2020). In practice, the amended act shortened the integration period from three and a half years to three years and shifted the responsibility of choosing and paying for integration training and exams from municipalities to individual immigrants. Individuals could now choose either to self-study or to join an integration course at approved schools with a social loan of up to EUR 10,000 (DUO, 2022).[6] The Service Execution Education (*Dienst Uitvoering Onderwijs*: DUO) administers the exam and centrally approves all integration courses. Hence, in comparison with *Wi*2007 where the integration process heavily depended on municipalities, *Wi*2013 reduces the heterogeneity in available information and the quality of integration courses across municipalities. The detailed changes to the integration requirements are presented in Table B.1 with a summary provided in Figure 2.1.

[5] For example, which modules to choose and whether to take the basic language exam or a higher level of language exam (NT2).

[6] The estimated cost of the integration course is EUR 7,000 (EUR 14/h x 500h). According to Fiolet Taaltrainingen (nd), the average preparation time is 500 hours. In addition, the average cost of approved integration courses for Dutch beginners in Amsterdam is EUR 13.8 per hour (Stichting Blik op Werk, nd).

Figure 2.1: Changes in integration requirements in the Netherlands



Note: Based on the information by Dutch Central Government (nd); DUO (nd); Algemene Rekenkamer (2017); IND (nd); Department of Dutch, UCL (nd), this chart is visualised by the authors.

The act of *Wi*2013 specifies the integration requirements that all refugees in our sample needed to fulfil. Under the *Wi*2013 the integration exam consisted of five components: listening; reading; writing; speaking Dutch; and knowledge of Dutch society (*Kennis van de Nederlandse Maatschappij*: KNM).[7]

On passing the integration exam, individuals will receive an integration diploma and their social loan will be waived as long as the exam was completed within the given integration period. Individuals who fail the exam can request to extend the integration period if they have grounds to do so (e.g. if they experienced a delay due to a backlog in processing civic exams or there is a long waiting time for social housing). If DUO accepts the request, refugees are given an additional integration period without penalties. If not, they will have to pay a fine

[7] Passing the Dutch language component requires individuals to achieve a minimum level of A2 (basic). However, immigrants can choose to sit the NT2 Dutch exam provided by the Dutch government which is at a higher proficiency level (B1 or B2: intermediate).

of up to EUR 1,250 and will not receive a renewed temporary asylum residence permit after the initial residence permit expires (Algemene Rekenkamer, 2017).[8] Passing the integration exam is, therefore, compulsory for all refugees over 18 years old if they wish to remain in the Netherlands.[9]

2.2.2 Orientation about the Dutch Labour Market (ONA)

In January 2015 a new component called Orientation about the Dutch Labour Market (*Oriëntatie op de Nederlandse Arbeidsmarkt*: ONA) was added to the existing integration requirements. This new component was publicly announced in the Official Gazette on 31 October 2014. The ONA is based on the idea that a lack of understanding of how the labour market works in the Netherlands makes it difficult for foreign-born workers to find suitable jobs (Broom NT2, nd). The ONA was therefore designed to deepen immigrants' understanding of the labour market through working on a portfolio of assignments alongside a mock interview or an additional 64-hour course.[10] Immigrants taking the ONA component are required to research the types of jobs they are interested in, to identify the qualifications or skills that are needed for these specific jobs, to prepare a CV to apply for their desired job,

[8] Usually, refugees are issued a Temporary Asylum Residence Permit which is valid for five years and which needs to be renewed if they hope to stay longer.

[9] Unless they meet exemption requirements such as providing evidence of medical difficulties or obtaining a diploma taught in Dutch.

[10] The duration to complete the ONA varies by individual and the courses that the individual takes, but it typically takes approximately 4-5 months to complete the ONA based on a 12-week course delivered by the University of Groningen (Language Centre, University of Groningen, nd) and a 6-week waiting time for the final interview (DUO, nd).

and to sit mock job interviews (DUO, nd). Hence, the ONA is expected to improve the labour market performance of immigrants and to specifically help them find and secure a job that matches their skills and/or qualifications. According to DUO, refugees tend to take the ONA as the last component within their integration exams, after developing some proficiency in the Dutch language.

The requirement to complete the ONA component came into effect in January 2015 and applies to asylum seekers who were granted an asylum residence permit after 1 January 2015. Refugees who obtained the residence permit before that date are exempt from the ONA component, even if they attempt the integration exam after 2015. The completion of the ONA component, therefore, depends entirely on the date of the asylum residence permit and is not a decision of individual refugees.[11] This precise cut-off date due to the policy change, therefore, provides us with a clear identification mechanism to examine the effect of this labour market training component.

As Table 2.1 shows, there are only small differences in the integration exam pass rate or the duration of the integration period before and after the introduction of the ONA. Under *Wi*2013, the overwhelming majority of refugees fulfilled the integration exam requirement with an average duration of 31-33 months.

[11] We later show that our results are robust to the possibility that refugees could have influenced their timing of arrival in order to avoid the ONA

Table 2.1: Pass rate and duration of integration for asylum migrants by cohort

	2013	2014	2015	2016
Pass Rate	98.3%	98.2%	97.0%	95.3%
Average integration period (months)	33	33	31	32

Note: The results of integration exams by cohorts as of 1 May 2022 are shown (DUO, 2022). 'Pass' refers to those who completed the civic integration exam/NT2 exam and the ONA for Cohort 2015+, including those who were accepted for exemption due to their health condition or diploma in Dutch. The rate shows the share of those who fulfilled all the integration requirements or were exempted, out of the total number of refugees who were obliged to fulfil the integration requirements in the same year. Average integration period refers to the average time in months to fulfil the integration obligation, excluding delays due to the 'Special circumstances' accepted by DUO.

2.3 Data

We base our analysis on two Dutch administrative datasets provided by Statistics Netherlands (CBS). Our main dataset is *Asielcohort* which contains detailed demographic and administrative information on the universe of refugees who obtained an asylum residence permit in the Netherlands since 2014. *Asielcohort* provides, for instance, information on the country of origin, gender, education in the Netherlands, integration exam status, and the date the asylum residence permit was issued.

A key feature of the Dutch government's refugee settlement is the quasi-random allocation of refugees to social housing locations. Applicants are initially allocated to refugee centres across the Netherlands where they stay until they know whether they are granted asylum. They are then moved into social housing in municipalities. In the absence of exceptional circumstances, refugees have no say in the locations to which they are allocated and such allocations are made in proportion to the population levels of each municipality.[12] In our sample, approximately 90% of the refugees have remained in their initial assigned location without relocating throughout the study period.

We complement the *Asielcohort* data with information on labour market performance at the individual level from *SPOLISBUS*, which provides monthly data on the employment status, earnings, and hours worked for the population of employed individuals in the Netherlands. Additionally, the dataset includes a unique job identifier for each worker in each firm, which consists of a combination of the job's start and end dates, the firm ID, and the employee ID. This unique identifier allows for accurate tracking and analysis of individual job histories within the study period. Using the unique social security number of each individual (employee ID), we merge these two datasets for the period from 2014 to 2021. Finally, for the firm level analysis, we link these two datasets to *Production* statistics, which is a firm register and encompasses almost the universe of firms in the Netherlands. Production statistics include detailed information on firms, for example, gross value added, wage bill, personnel costs etc. The final data creates a linked employer-employee dataset for 8 years.

These detailed datasets allow us to distinguish which exam the individual took based on

[12] Exceptional circumstances include the need for medical treatment or the presence of immediate family members within a particular location. However, even if these criteria are met, in practice it is often impossible for the Dutch authorities to meet refugees' needs or preferences due to long waiting lists for social housing. Furthermore, refusing social housing due to its location would result in a refugee remaining in the refugee centres (*AZCs*), where again they are allocated randomly.

the date of their asylum residence permit. A limitation of our data is the lack of information on human capital obtained prior to arrival to the Netherlands, such as educational attainment, employment experience or occupations in the home country. We overcome this limitation by controlling for individual level fixed effects in our main empirical specification.

2.4 Empirical Methodology

In order to evaluate the impact of the ONA component on labour market outcomes, we compare trajectories in labour market performance since the beginning of the integration period of those who completed the ONA as part of the new civic integration exam and those who passed the old exam without the ONA. To deal with potential biases caused by differences in the observable characteristics of the two groups, we construct a matched sample using PSM techniques.

2.4.1 Study Design

To assess the impact of the ONA component on the economic integration of refugees, we compare the labour market performance between those who passed the new exam with the ONA component (the treated) and a control group of refugees who were required to pass a standard integration exam without the ONA. The treatment group is defined as refugees who were granted an asylum residence permit after 1 January 2015 and are thus required to pass the new integration exam with the ONA component. The control group consists of refugees who were granted an asylum residence permit in 2014 and were only required to complete the five components in the standard civic integration exam without the ONA. In other words; the treatment is those who obtained the asylum residence permit after the cutoff day and expressed in the equation below:

$$ONA_i = 1\{x_i \geq c\} \tag{2.1}$$

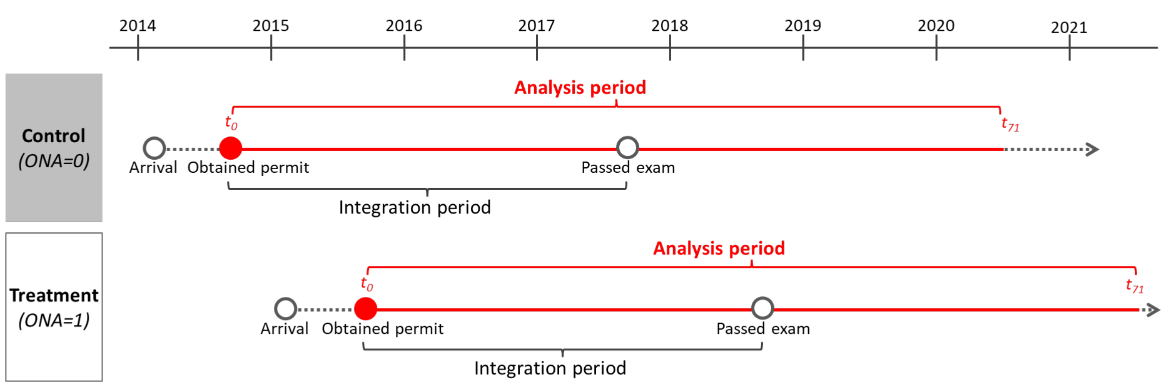
where x_i refers to the date the individual i obtained the residence permit while c shows the cutoff date of ONA introduction on 1 January 2015. $ONA = 1$ for the treatment group and $ONA = 0$ for the control group.

We restrict our data to the 2014 and 2015 cohorts of *Asielcohort* to ensure that the treated and control groups are sufficiently similar in terms of observable individual-level characteristics such as country of origin. Moreover, we limit our sample to refugees from these two cohorts who have successfully completed the civic integration exam, and who are of working age (i.e. 18 to 65).[13] After the application of these selection criteria, our sample includes 13,501 unique individuals, and approximately one million observations.

[13] These restrictions are made to try to ensure comparability between and within our treatment and control groups. We also exclude other 'non-standard' cases e.g. we omit those who passed a higher level language exam (at B1 or B2 level) and those who left the Netherlands. As a robustness check, we also include those who left the Netherlands (Figure B.3 in Appendix).

Our analysis period extends for eight years from the month the individual obtained the asylum resident permit. Once residence permits are issued individuals are allowed to work without restriction and are obliged to fulfil the civic integration requirements within three years, in other words, the three-year integration period begins at t_0. Figure 2.2 outlines the study period for our analysis and illustrates the date of obtaining the asylum residence permit and the date of passing the exam for two hypothetical individuals, one in the treatment group and one in the control group. For the control group (i.e. the 2014 cohort), we have observations until 2020 while we follow the treatment group until 2021 to ensure an equivalent post-permit period of analysis.

Figure 2.2: Standard time frame for control and treatment groups



Note: This chart shows the integration time scale for treatment and control groups. t_0 refers to the month that individuals obtained their residence permits and thus started their integration periods. Individuals whose asylum residence permits were granted in 2014 took the old exam and constitute the control group. Those whose permits were granted in 2015 took the new exam, containing the ONA, and constitute the treatment group.

2.4.2 Summary Statistics

Table 2.2 and Table 2.3 present the summary statistics of our sample at the individual and at firm level, respectively. The summary statistics are broken down by the *treatment* and the *control* groups in Column (1) and (2). While Table 2.2 suggests that the average arrival age for both groups was around 30 years old and that both groups face extremely low employment probabilities of 1 to 2% at t_0, there are significant differences between the two groups in terms of gender and country of origin. The share of female refugees in the treatment group is almost 10 percentage points higher than that of the control group (39.86% and 29.47%, respectively), and Syrians account for more than two-thirds of the treated compared with 57.79% of the control group. Bevelander (2020) points out that female refugees lag behind male refugees in terms of employment outcomes and although all refugees tend to perform worse in the labour market in comparison to immigrants from the same region, this gap is significantly larger for refugees from the Middle East and Africa (Fasani et al., 2022). Helgesson et al. (2019) also indicate that refugees from Syria and Somalia in particular experience a greater risk of long-term unemployment. Table 2.2 also indicates that the size of firms that treated refugees in employment work for tends to be larger with a slightly lower level of labour share.

To mitigate the impact of the over-representation of women and Syrians in the treated group, we select a control group of refugees with similar characteristics to the treated group, using a PSM method. We control for individual characteristics, including gender and country of origin, as well as the duration of the integration period as a proxy for the ability and/or

Table 2.2: Summary statistics - Individual characteristics

	Before matching				After matching			
	(1) Treatment		(2) Control		(3) Treatment		(4) Control	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Panel A: Time-invariant variables								
Asylum claim (months)	2.72	3.28	2.91	2.49	2.72	3.28	2.71	2.63
Integration period (months)	29.04	5.00	33.62	7.64	29.04	5.00	29.11	5.34
Arrival age	30.60	8.59	29.45	8.42	30.60	8.59	30.85	8.93
Female	39.86%		29.47%		39.86%		39.46%	
Country of origin								
Syria	70.90%		57.79%		70.90%		71.94%	
Eritrea	16.27%		25.88%		16.27%		14.17%	
Iran	1.71%		2.23%		1.72%		1.84%	
Family size	1.48	1.20	1.22	0.86	1.48	1.31	1.31	0.99
Panel B: Time-variant variables at t_0								
Education in the Netherlands								
No education	98.16%		98.06%		98.16%		98.33%	
Low/Middle	1.67%		1.69%		1.67%		2.11%	
High	0.17%		0.25%		0.17%		0.57%	
labor market performance								
Employment prob.	0.02	0.13	0.01	0.10	0.02	0.13	0.01	0.10
Hours worked	58.90	55.16	64.62	47.60	58.90	55.16	60.00	49.45
Hourly wages	10.10	2.83	9.79	2.53	10.10	28.23	10.29	2.69
Individuals	7,173		6,328		7,172		6,328	

Note: Time-variant variables are reported in the month that each individual obtained the asylum residence permit (t_0). Family size refers to the number of people who arrived in the Netherlands together (family size upon arrival). labour share is defined as the total wage bill divided by value added. Hours worked and hourly wages are conditional on being employed.

Table 2.3: Summary statistics - Firm characteristics and mobility (whole period)

	Before matching				After matching			
	(1) Treatment		(2) Control		(3) Treatment		(4) Control	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Value added (in €1000)	4,919.72	3,932.07	4,771.61	3,874.50	4,919.72	3,932.07	4,816.01	3,956.80
Labour share	0.56	0.18	0.57	0.17	0.56	0.18	0.56	0.17
Number of moves	0.22	0.44	0.49	0.55	0.22	0.44	0.54	
Observations	469,521		469,535		469,456		469,535	

Note: Labour share is defined as the total wage bill divided by value added. Firm characteristics are conditional on being employed.

motivation to adjust to the Dutch labour market. We perform the matching on the basis of a Gaussian Kernel approach with a bandwidth of 0.01 and common support.[14]

Table 2.2 also shows the summary statistics after matching in Columns (3) and (4). After matching, the treatment and control groups contain similar proportions of females and Syrians. Table B.2 in the Appendix reports the results of the matching balancing tests which show no statistically significant differences across the two groups.

2.4.3 Econometric model

The matched sample contains two asylum cohorts with similar observable individual characteristics except for whether or not they have taken the ONA. We compare the trajectories of labour market outcomes between the treatment and control to assess the impact of the ONA on the labour market performance of refugees. Our baseline specification is as follows:

$$\begin{aligned} Y_{irt} = \alpha + \sum_{q=1}^{23} \beta Period_{iq} + \sum_{q=1}^{23} \gamma (Period_{iq} * ONA_i) + \rho Z_{it} + \mu_i \\ + \tau_t + \sigma_r + \upsilon_{it} \end{aligned} \tag{2.2}$$

[14] Kernel matching maximises precision by retaining all matched controls and assigning greater weights to better matches since this approach selects individuals from the control group, who are similar to the treated, and assigns weights to each control based on the propensity score distance (Garrido et al., 2014). Further details of the matching specification are described in Appendix B.2.

where Y_{irt} refers to the labour market performance (i.e. the probability of being employed, the logged number of hours worked or logged hourly wages) of individual i in municipality r at time (month) t. We consider the probability of being in employment as a proxy for labour market participation, hours worked as an indicator of employment quantity, and hourly wages as a measure of job quality. $Period$ indicates dummies for each quarterly time period, q, since the beginning of the integration period[15], Z refers to time-varying individual controls such as education in the Netherlands while μ is individual fixed effects (FE) to capture time-invariant individual characteristics such as gender, country of origin, education attainment and working experiences before arrival in the Netherlands. τ is interactions of time FEs with country of origin FEs (Year-Quarter-Country: Y-Q-C) to capture seasonal effects and the possibility of changing attitudes towards refugees from certain countries or the introduction of policies that affect refugees from some countries more than others. σ are residence municipality FEs that controls for structural differences between locations. The standard errors are clustered at the municipality level, given that access to integration courses, labour market conditions and industrial structures differ by municipality.

As part of our robustness analysis we undertake a complementary Regression Discontinuity (RD) approach following Foged et al. (2022) who apply this method in a similar setting to ours. RD exploits the fact that the ONA has a sharp implementation date (cut-off) and relies on the likelihood that the characteristics of refugees will be comparable either side of

[15] As the time period is here defined in quarters, $Period_0$ includes t_0 to t_2 while $Period_{23}$ is equivalent to t_{69} to t_{71}. The reference period is $Period_0$.

the cut-off. Individual FEs are therefore not included. Our results using this approach are comparable to those based on equation 2.2. These results and the specification details are presented in Appendix B.6.1.

Referring back to equation 2.2, the coefficient γ captures the difference in the outcome for the treatment group over and above those who did not take the ONA, and is our coefficient of particular interest. Coefficient β shows the common trajectories across treated and control groups and captures the effect of the five components of the initial integration exam on labour market outcomes as well as the effect of time spent in the Netherlands.[16] Since all components of the integration exam support the development of Dutch language and social skills, we expect β and γ to be significantly positive. Two of our outcome variables; hours worked and hourly wage, are only observed when an individual is employed. When we consider these outcome variables, therefore, we restrict our sample to those who are employed at time t. This restriction results in a significant reduction in the sample size for the specifications with these two outcome variables.

Finally, to assess whether there is heterogeneity in the impacts of the ONA, we replicate this approach on a selection of sub-sample groups, namely refugees from Syria; those from

[16] Since the dummies of $Period$ capture gaps in labour market outcomes for each specific period, we do not include the number of months since migration. As Albert et al. (2021) indicate, wage gaps between immigrants and natives converge as immigrants stay in the host country longer due to the cultivation of locally applicable human capital and the adjustment of unobserved characteristics.

other countries; males; and females.[17]

2.5 Estimation Results

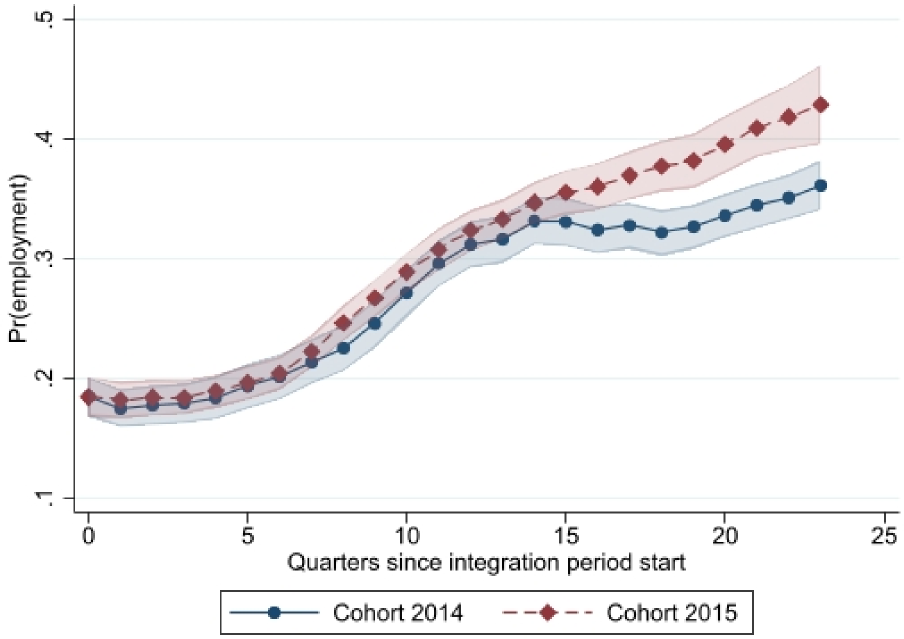
2.5.1 Baseline specifications

Figure 2.3 presents the results of our baseline estimation of Equation (2.2) for the probability of being employed. We observe that both cohorts experience an increase in the probability of being employed approximately a year and a half after the start of their integration period. However, the difference in trajectories between the two cohorts becomes statistically significant only after approximately 16 quarters (4 years) and from there it increases over time. By the end of the analysis period (quarter 23), those in the treatment had employment probabilities of approximately 42%, compared to 36% for the control, a difference of 6.4 percentage points. Since the average duration to complete the integration exams including the ONA is approximately 3 years, the timing of the divergence of the cohorts is consistent with the ONA having a positive impact on the treatment group.

The results presented in Figure 2.4 indicate that the other labour market outcomes also improved throughout the period. In these estimations sample size decreases by 26% as they

[17] To identify the matched treated and controls and the weights within each sub-sample, we replicate the propensity score matching procedure described in Appendix B.2 with the same set of variables as in the full sample except for the variable that specifies the sub-sample. For instance, the matching equation for the female sub-sample omits the gender variable. All the matching estimations use Kernel matching with a 0.01 bandwidth.

Figure 2.3: Baseline estimation
(*Employment probability)*



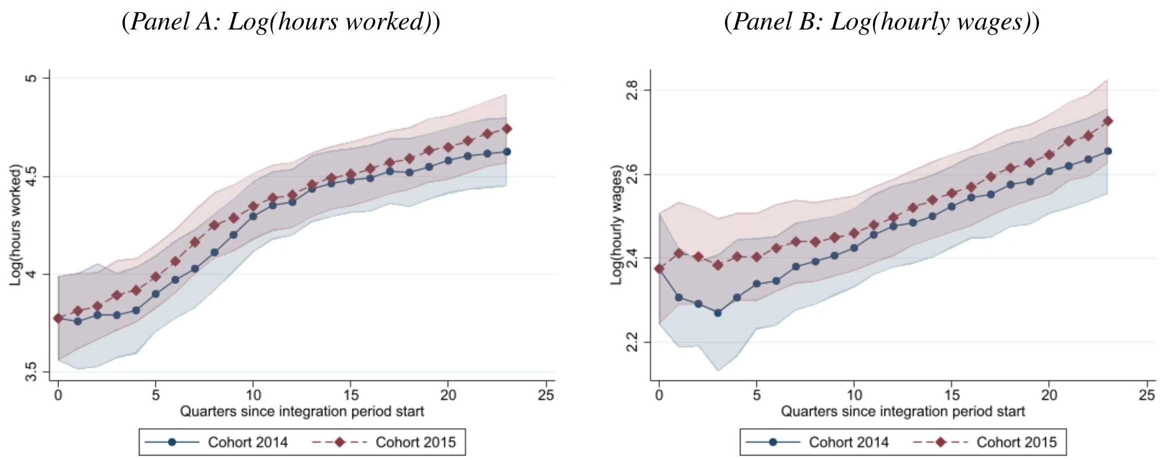
Note: The shaded regions show the 95% confidence intervals. Cohort 2014, who obtained the asylum residence permit in 2014, were not required to complete ONA whereas those who received the permit in 2015 (i.e. Cohort 2015) were obliged to take the ONA. The employment probability of Cohort 2014 is calculated by summing the constant (i.e. α) and the coefficient of the period dummies of the period (i.e. β) from Equation (2.2). For Cohort 2015, it is computed with α, β and the coefficients of the interaction (i.e. γ). Estimations include individual, year-quarter-country, and municipal fixed effects as well as individual time-varying characteristics. The standard errors are clustered at the municipality level.

are based on refugees who were in employment at any point within the study period. *Panel A* shows that both cohorts work 0.85 log points hours longer per month by $Period_{23}$ than at the beginning of the integration period. This means that over this period hours worked increased by a factor of 2.3.[18] *Panel B* also shows a rise in hourly wages for both cohorts by 0.281 log points, translating into a 30% increase in hourly wages. These effects are economically significant, and the improvement of labour market performance for both cohorts throughout

[18] We estimate average hours worked of both treatment and control increased from around 40 hours per month to 110 hours per month after taking the civic integration exam.

the period is consistent with the findings of Albert et al. (2021). Our results therefore suggest that the completion of the integration exam and the duration of stay in the Netherlands increase the acquisition of locally applicable knowledge by refugees. However, we do not find any statistically significant differences between the cohorts.

Figure 2.4: Baseline estimations



Note: The shaded regions show the 95% confidence intervals. Hours worked and hourly wages are conditional on being employed. Cohort 2014, who obtained the asylum residence permit in 2014, were not required to complete the ONA whereas those who received the permit in 2015 (i.e. Cohort 2015) were obliged to take the ONA. Cohort 2014 is calculated by summing the constant (i.e. α) and the coefficient of the period dummies of the period (i.e. β) from Equation (2.2). For Cohort 2015, it is computed with α, β and the coefficients of the interaction (i.e. γ). Estimations include individual, year-quarter-country, and municipal fixed effects as well as individual time-varying characteristics. The standard errors are clustered at the municipality level.

These results suggest that, in contrast to general linguistic training, (*e.g.* (Lochmann et al., 2019) specifically targeted labour market participation training makes a positive difference in employment outcomes of refugees. However, the ONA effects only the likelihood of participating in the labour market and is not associated with longer hours of participation or with better (more highly paid) jobs for those who are employed. This finding reflects the fact that the ONA component focuses on the skills required to search and apply for jobs that match the knowledge of the applicants. It does not target improvements in human capital, such as advanced language proficiency or more complex-skills training, that is required to access higher-skill occupations, such as those that are more communication-intensive (Algemene Rekenkamer, 2017; Foged et al., 2022). In other words, this result implies that the ONA component boosts employment by improving the knowledge required to enter the labour market, but does not lead to an upgrading of the skills that are needed for an improvement in the quantity or the quality of employment.

2.5.2 Potential endogeneity relating to the timing of the arrival

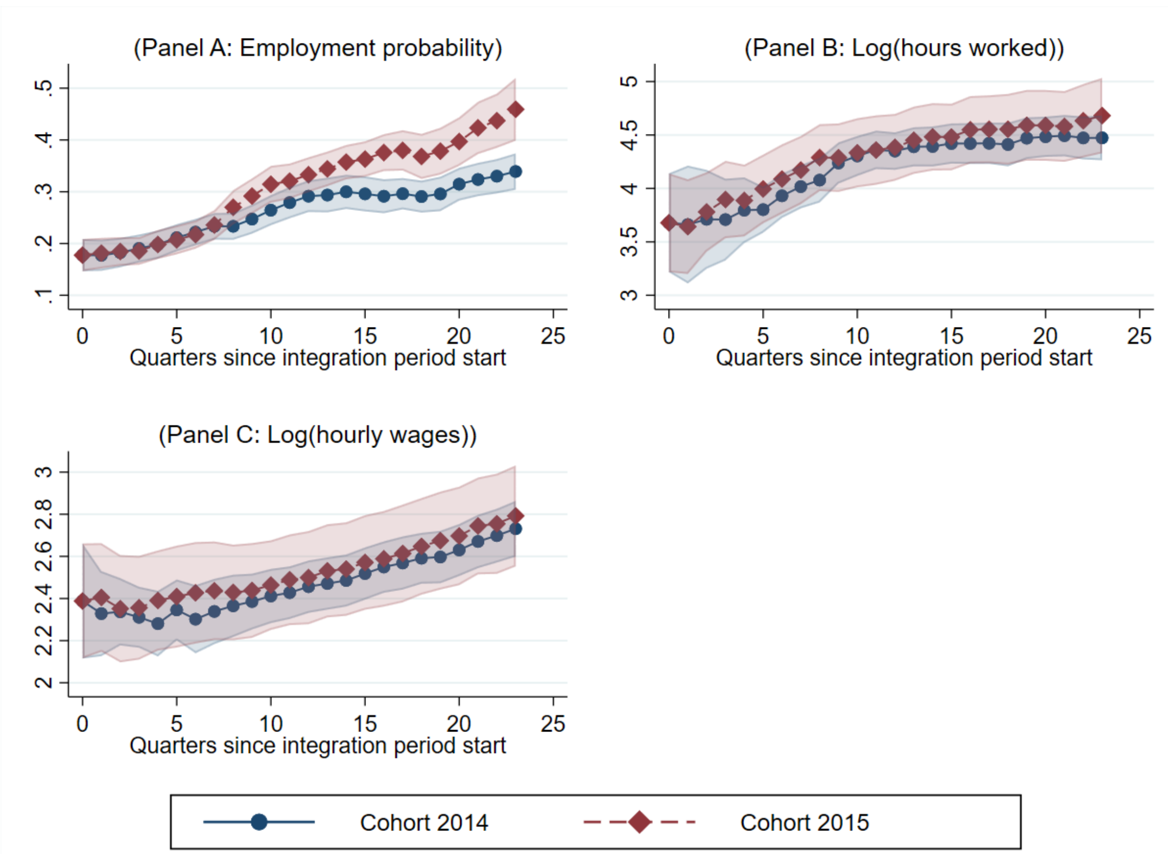
A potential endogeneity concern could arise if asylum seekers intentionally changed their date of arrival in the Netherlands as a result of the implementation of the ONA. For instance, less able asylum seekers may have chosen to arrive prior to January 2015 in order to avoid the ONA, resulting in an upward bias in our estimation of the employment benefits of the ONA. While such a scenario seems unlikely, as studies tend to show that refugees have a limited knowledge of migration policies within and between countries (Robinson and Segrott, 2002; Crawley and Hagen-Zanker, 2019) and there are many factors that refugees cannot influence *e.g.* administrative processes, it is something we now explore.

To do so, we focus on the fact that the introduction of the ONA was only officially announced on 31 October 2014.[19] We, therefore, restrict our sample to those who arrived in the Netherlands between June and October 2014 on the basis that nobody in this restricted sample would have been aware of this policy change before their arrival in the Netherlands. This restricted sample still provides treatment and control groups since it is the date on which individuals were awarded their asylum permits that determines exposure to the treatment and there was generally a gap of some months between arrival and being awarded a permit. Hence, anyone who arrived between June and October 2014 but whose permit was awarded in January 2015 or later would form our treatment group. The restricted sample consists of approximately 30% of the total sample.

Figure 2.5 provides the results of our baseline estimations using the restricted sample. *Panel A* indicates that employment probabilities of both cohorts increase from around 18% at Period 0 but diverge in a statistically significant manner after around 15 quarters. By the end of the sample period the 2014 cohort have an employment probability of approximately 35%, compared to around 45% for the 2015 cohort. *Panels B* and *C* indicate that there are again no significant differences between the two cohorts in terms of hours worked and hourly wages. These results are consistent with our main findings and confirm that these are not unduly biased by endogeneity concerns due to refugees potentially bringing forward their arrival time.

[19] The Official Gazette 2014-404, which announced the amendment of the Civic Integration Regulation regarding the addition of a practical exam for orientation on the Dutch labour market, was published 31 October 2014.

Figure 2.5: Baseline estimations with restricted sample - cohort arrived June-Oct 2014



*Note:*The sample is restricted to those who arrived in the Netherlands between June and October 2014. Cohort 2014, who obtained the asylum residence permit in 2014, were not required to complete the ONA whereas those who received the permit in 2015 (i.e. Cohort 2015) were obliged to take the ONA. The employment probability of Cohort 2014 is calculated by summing the constant (i.e. α) and the coefficient of the period dummies of the period (i.e. β) from Equation (2.2). For Cohort 2015, it is computed with α, β and the coefficients of the interaction (i.e. γ). Estimations include individual, year-quarter-country, and municipal fixed effects as well as individual time-varying characteristics. The standard errors are clustered at the municipality level.

2.5.3 Heterogeneity analysis

2.5.4 Refugee heterogeneity

As indicated in Table 2.2, a large proportion of the refugees in our sample originate from Syria. The literature on the recent waves of refugees indicates that Syrian refugees are typically characterised by higher educational attainment and are more likely to suffer from health problems (Chung et al., 2018; Aksoy and Poutvaara, 2021). Since both educational attainments and health conditions heavily affect labour market outcomes, the impact of the civic integration exam on the labour market performance of Syrian refugees may differ from that for refugees from other countries. While we are able to control for physical and mental health difficulties, education obtained in the Netherlands and individual FEs, we nevertheless here explore whether our findings differ by country of origin. Specifically, we estimate our baseline specification on two sub-sample groups: refugees from Syria; and refugees from countries other than Syria. The literature on refugees also highlights the limited labour market participation of female refugees in comparison to males (Bevelander, 2020). Indeed Syrians and females in our sample have labour market participation rates that are approximately 50% lower than those of other nationalities and males.[20] Therefore, we also explore heterogeneity across gender by splitting our sample into males and females.

Figure 2.6 presents the results for employment probability to assess the heterogeneity in

[20] Summary statistics of the matched samples for Syrians and females are available upon request.

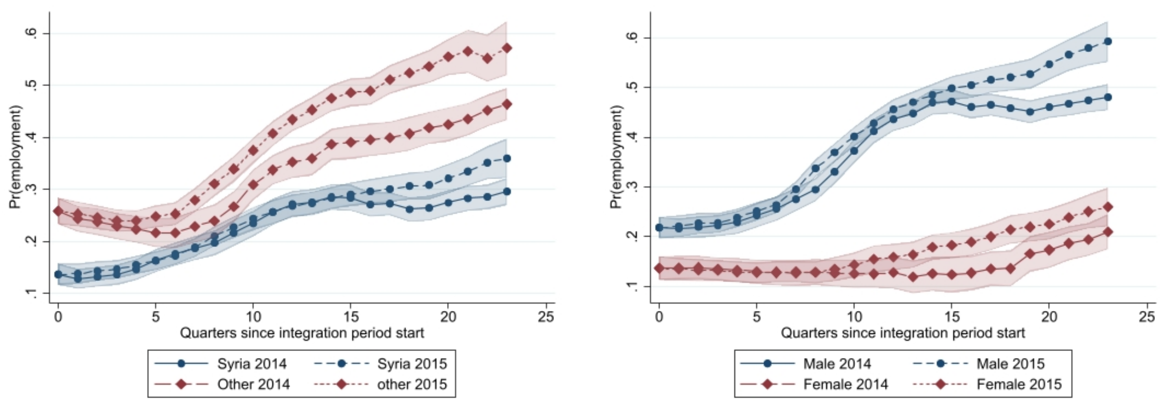
the effect of the ONA by country of origin and gender.[21] *Panel A* considers the differences between Syrians and other nationalities, while *Panel B* focuses on gender differences. Figure 2.6 shows that Syrian refugees (*Panel A*) and female refugees (*Panel B*) under-perform in the labour market in comparison with refugees from other nationalities and with male refugees, respectively. *Panel A* shows that the main beneficiaries of the ONA appear to be refugees from countries other than Syria for whom the treatment group displays a statistically significant difference in employment probability from around 8 quarters (2 years). In contrast, for Syrians, the difference in the treatment and control groups arises only after around 16 quarters and this difference is barely statistically significant. *Panel B* presents a similar pattern and shows that while we observe a divergence between the trajectories of treated and controls for both male and female sub-samples, this divergence is only consistently statistically significant for males. Compared to the difference of 6.4 percentage point between the treatment and control groups in the baseline estimates, we show that the effect of the ONA on males is approximately 11%.

These results echo previous findings that females and refugees from Syria are more likely to struggle to be integrated into the labour market than males or refugees from other countries(Helgesson et al., 2019; Bevelander, 2020). According to Albrecht et al. (2021), the low labour market integration of female refugees can be explained by their limited work experience and lower educational attainment in comparison to male refugees. Moreover and

[21] We focus only on employment probability due to reasons of space.

Figure 2.6: eterogeneity: Employment probability

(Panel A: Country of origin) (*Panel B: Gender*)



Note: Cohort 2014, who obtained the asylum residence permit in 2014, were not required to complete ONA whereas those who got the permit in 2015 (i.e. Cohort 2015) were obliged to ONA. The employment probability of Cohort 2014 is calculated by summing the constant (i.e. α) and the coefficient of the period dummies of the period (i.e. β) from Equation (2.2). For Cohort 2015, it is computed with α, β and the coefficients of the interaction (i.e. γ). Estimations include individual, year-quarter-country, and municipal fixed effects as well as individual time-varying characteristics. The standard errors are clustered at the municipality level.

as pointed out by Helgesson et al. (2019) and Fadhlia et al. (2022), the under-performance of Syrian refugees in the labour market can be explained by there being a disproportional share of that group who suffer from mental health issues. At the same time, refugees from Syria, especially females, are more likely to be highly educated (Aksoy and Poutvaara, 2021) but may face challenges accessing high-skill occupations because of the limited applicability of their human capital in the context of the Netherlands. In our sample, 75% of female refugees are from Syria (in comparison to 70% of males), where female labour market participation is as low as 16% (World Bank, 2022). This statistic suggests that female refugees in the Netherlands may have limited work experience which forms an obstacle to entering the labour market.

Moreover, it is notable that the ONA does not help to close the gap in employment probabilities between males and females and between Syrians and non-Syrians. In fact, the opposite is true. In Cohort 2015, the male-female employment probability gap was 33 percentage points by the end of the sample period compared to 27 percentage points in Cohort 2014. Similarly, the Syrian-Non-Syrian employment probability gap was 21 percentage points in Cohort 2015 compared to 17 percentage points in Cohort 2014.

2.5.5 Firm heterogeneity

A potential limitation of previous studies is that the increased refugee inflows in Europe may result in some firms employing more refugees regardless of whether those refugees have benefited from a labour market policy such as the ONA. If such an effect operates regardless of the ONA, not controlling for firms' unobserved demand for, and policies aimed at refugee employment could bias the estimates. In this section, we, therefore, ascertain whether the effect of the ONA is sensitive to unobserved firm effects. We also explore whether passing the ONA component influences the types of firms that refugees are likely to work for. Since the ONA is intended to increase the ability of refugees to find suitable employment that matches their skills and abilities, we may observe treated refugees being more likely to work for more productive, higher-skill firms that are typically larger in size and offer higher wages.[22] The literature points out that given difficulties to access jobs that match refugees' skills and

[22] Although the conventional wisdom is that high value-added firms tend to employ more skilled and more productive workers and pay higher wages, this may not always be the case, e.g. Lochner and Schulz (2021).

education, they are more likely to be employed in low-skilled services sectors and/or high labor-intensive firms (Fasani et al., 2022). This implies that refugees with the ONA may be less likely to work for such firms compared to refugees in our control group.

As a first step, we augment the baseline model of hours worked and hourly wages by including firm fixed effects, denoted by κ_f in our estimations as below:

$$Y_{ifrt} = \alpha + \sum_{q=1}^{23} \beta Period_{iq} + \sum_{q=1}^{23} \gamma (Period_{iq} * ONA_i) + \rho Z_{it} + \mu_i \\ + \kappa_f + \tau_t + \sigma_r + \upsilon_{it} \tag{2.3}$$

We then re-estimate our baseline equation conditional on being employed, but change the dependent variable, Y_{ifrt} to firm size, to capture the probability of a refugee working for a large firm, where we define firm size in terms of total value added. To examine whether passing the ONA makes refugees more likely to work for firms with a low labour intensity we replace the dependent variable with labour share which is defined as the total wage bill as a share of total value added. For both variables, the median value of value added and labour share is taken as the threshold to define large firms and low labour share firms, respectively, and is calculated using data on all firms across the Dutch economy.[23]

[23] To explore firm heterogeneity effects further Section 2.5.6 examines differences in firms' wage levels.

We firstly find that our results are unaffected by the inclusion of firm fixed effects i.e. we continue to find no statistical difference between the treatment and control for our measures of employment outcomes hourly wages and hours worked, suggesting that unobserved firm heterogeneity is not unduly influencing these results.[24]

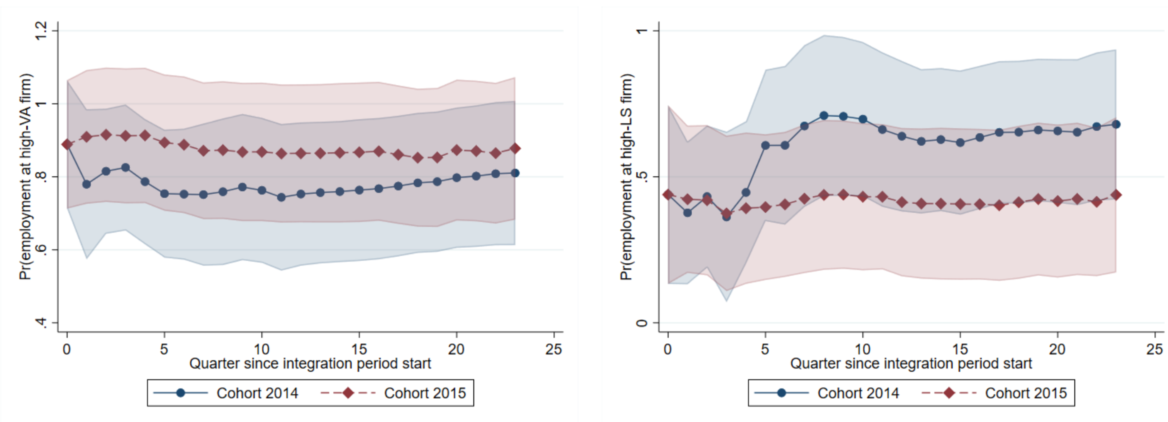
Turning to our estimations of firm size and labour share, Figure 2.7 provides our results. In terms of firm size, Panel A shows that treated refugees are more likely to work for large firms than refugees in the control group although the difference between the two is not statistically significant. Panel B illustrates the results for labour share. As expected, we see that refugees who've passed the ONA are more likely to work for firms with low labour share, although again the differences between the treatment and the control groups are not statistically significant.

These results therefore provide some limited support for the notion that the ONA improves refugees' abilities to find employment that matches their skills resulting in them being more likely to be employed in larger, higher-skill, lower labour share firms.

[24] These results are available upon request. Firm fixed effects are only included in the estimations that are conditional upon being employed.

Figure 2.7: Firm heterogeneity

(Panel A: Employment in high value-added firms) *(Panel B: Employment in high labour share firms)*



Note: High value-added firms are defined as firms with value-added above the median of EUR 1,285 million, and high labour share is above the median 0.568.

2.5.6 Potential mechanisms

We now turn to the potential mechanisms through which the ONA's labour market training could improve labour market participation. The first and most obvious mechanism is if the training within the ONA genuinely improves the knowledge and abilities of refugees meaning they are better equipped for the labour market. We refer to this as a 'learning effect'. A second mechanism could simply be a signalling effect. If employers *perceive* individuals who passed the ONA component to be better trained and integrated into the workplace, they may be more likely to employ those individuals. Third, the ONA may increase labour market participation by stimulating refugee mobility. By providing the skills and confidence to seek work in more advanced labour markets, the ONA may generate greater movement to, for instance, urban areas where there is a greater volume and diversity of jobs available. Finally, as suggested by McKenzie and Rapoport (2010), there may also be a network effect. The network effect could arise if, having completed the training, individuals are better connected to Dutch citizens and fellow refugees, more confident about social interaction and better informed of employment opportunities. [25]

Learning effect

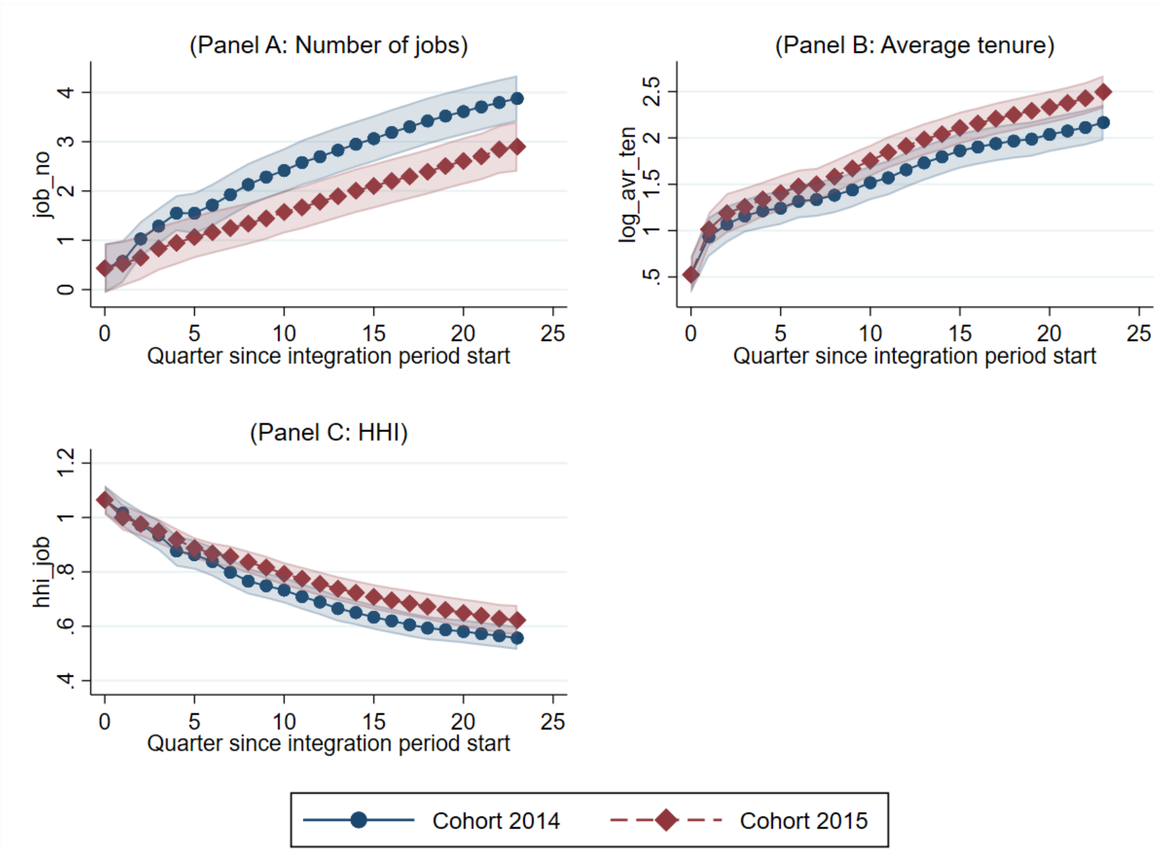
If the ONA affects employment probability via the learning mechanism then we may expect treated refugees to have greater job stability than those in the control group as a result of a better matching with employers. To test this hypothesis, we compare employment stability between the two cohorts. We replicate Equation (2.2) but replace the dependent variable with three different proxies for employment stability: i) the number of jobs; ii) the average tenure of jobs; and iii) the Herfindahl-Hirschman index (HHI) of jobs. This analysis is possible since our data allows us to observe a unique job identifier for each job-employer pair.[26] Accordingly, stable jobs will be associated with a lower number of jobs and a longer average tenure per job (where tenure indicates the number of months of employment in a specific job) and a lower value of HHI. HHI is a composite measure of the level of concentration applied to market dominance or diversity in labour forces (Rhoades, 1993; Hunt et al., 2015; Ozgen, 2021). In our context, greater job stability will result in a smaller number of jobs having a

[25] Networks are here defined in a broader sense, consistent with Lochmann et al. (2019), and not only in terms of same-country-of-origin networks examined by Dagnelie et al. (2019). With regards to the 'networking' mechanism, the introduction of the ONA component does not significantly increase the networking opportunities available to refugees. Refugees in both treatment and control groups will take the full civic integration exam, with the ONA component taken by the treatment group comprising only a small proportion of the total hours needed to meet the civic integration exam requirements. Therefore Cohort 2014 also has a significant opportunity to meet and interact with Dutch citizens and peer refugees while completing the integration courses. The additional hours associated with the ONA component would therefore appear to have little impact on the difference in networking opportunities between the two groups.

[26] All of these proxies used are conditional on being employed.

dominant share in the entire tenure of employment of each individual, which translates into a higher concentration index.[27]

Figure 2.8: Job stability



Note: The replication results of Equation (2.2) on job stability indicators, all of which are conditional on being employed. Average tenure by employers is shown in log terms. Estimations include individual, year-quarter-country, and municipal fixed effects as well as individual time-varying characteristics. The standard errors are clustered at the municipality level.

Figure 2.8 presents the results for our three measures of job stability. It indicates that Cohort 2015 have fewer jobs in comparison with Cohort 14 and this difference is significant from

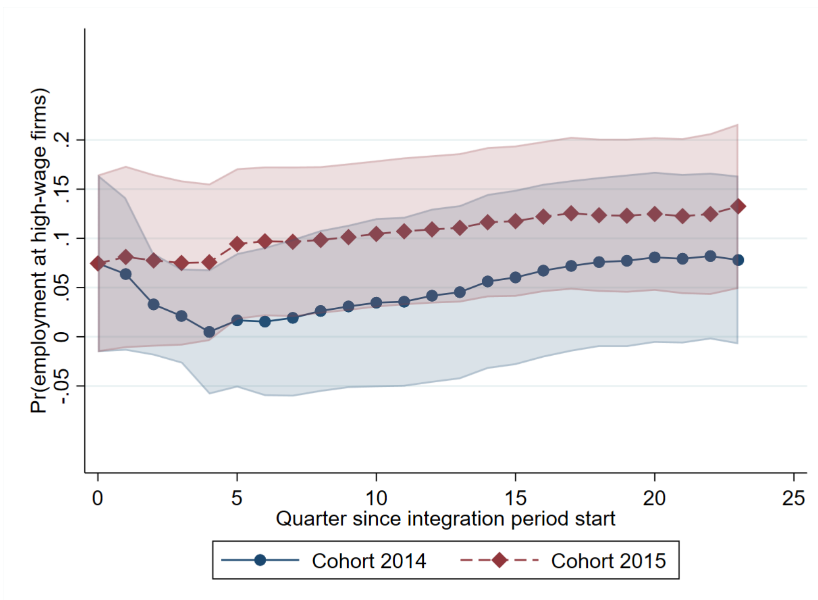
[27] The detailed descriptions of HHI, including the computation, are presented in Appendix B.5.

$Period_{11}$ onwards (*Panel A*). By the end of the period, refugees in Cohort 2015 have held, on average, 1 job less than those in Cohort 2014, where the mean number of job changes is 2.5. *Panel B* suggests that refugees in Cohort 2015 have, on average, 3 months longer tenure, and the difference becomes marginally significant from $Period_{19}$ onwards. Although the value of the HHI is larger for Cohort 15, the difference between the two cohorts is not statistically significant throughout the period (*Panel C*). Overall, these results imply that refugees in Cohort 2015 have more stable employment with significantly fewer job changes and a slightly longer tenure in each job. This suggests, compared to Arendt (2022)'s short term study, in the longer term in fact the policy helps refugees to find a more sustainable employment. In line with our expectations, these findings indicate the benefits of the ONA training to manifest themselves through better matching between refugees and employers and better orientation in the Dutch labour market.

Signalling effect

In terms of the second potential mechanism - the signalling effect - a training programme may signal to employers that those who have undertaken such training are harder working or more able in some unobserved dimension. However, this effect assumes that individuals are able to self-select into the training programme implying that higher productivity individuals will indeed opt-in while lower productivity individuals will opt-out(Lazear and Gibbs, 2014; Borjas, 2020). Where such self-selection is present, the completion of a training course or qualification allows employers to identify higher-productivity individuals. As mentioned earlier though, the ONA is compulsory from the 2015 cohort onwards and hence refugees are not able to choose whether or not to participate in it. In this setting, passing the ONA component does not reveal information on the unobservable ability of refugees and should not therefore act as a signal to employers.

Figure 2.9: Employment in high wage firms



Note: Based on Equation (2.2) with a high-wage firm dummy being the dependent variable. High-wage firms are defined as firms with hourly wages above the median of EUR 17. Estimations include individual, year-quarter-country, and municipal fixed effects as well as individual time-varying characteristics. The standard errors are clustered at the municipality level.

Nevertheless, there may be a pure 'sheepskin effect' whereby individuals who possess the ONA certificate are seen as being superior by employers in some way even if in reality they are not. A pure sheepskin effect could indeed result in greater employment probability, but is likely to result in a mismatch between worker ability and the skill-requirements of jobs, hence over-qualification, leading to shorter employment duration and a greater number of jobs over a fixed period of time. As we have discussed earlier, Figure 2.8 indicates that treated workers actually experience longer employment duration and fewer job changes, contrary to what we might see if a signalling effect were present.

To explore this further though we argue that high-wage firms are likely to be more receptive to credentialism implying such firms may be more inclined to employ individuals who have passed the ONA to ensure employees hold higher standards. Given we have already found that those with the ONA are not in fact paid more, if they are still more likely to be employed by high-wage firms we may argue that this can be seen as evidence of a signalling effect. To test this we repeat our baseline equation conditional on being employed, yet we change the dependent variable to the probability of being employed in the most productive firms, defined as those paying above median wages.[28] Figure 2.9 indicates that although passing the ONA leads to a higher probability of employment in more productive firms, the difference between the treatment and control group is not statistically significant. Therefore, individuals with the ONA are not paid more and are not more likely to be employed by high-wage firms. In sum, employers do not seem to use the ONA to distinguish between job seekers and do not perceive it to be a signal of higher ability or job performance.

[28] The mean and median pay level is rather similar in our dataset and changing the threshold from above 50^{th} percentile to 75^{th} percentile does not alter the results.

Geographic mobility towards urban areas

In many European countries refugees are commonly distributed to remote locations. However, little is known about whether completing mandatory integration training programmes provides refugees with the sorts of skills that influence the likelihood of them relocating to other, perhaps more urban, areas.

Such programmes are likely to generate two competing effects regarding the direction of mobility for refugees. On one hand, they may incentivise refugees to relocate to more urban areas, where a greater variety of job opportunities are available for individuals with diverse skill levels. On the other hand, labour market training can also assist refugees in securing employment within their initial assignment area, thereby encouraging them to remain there. In this section, we will discuss and explore further which of these effects prevail.

The advantages of agglomeration economies are most pronounced in urban areas and, in the case of refugees, are likely to be compounded by the presence of co-nationals which may enhance both the willingness and capacity for their economic integration as Chapter 1 suggests (Yumoto et al., 2023a). Eckert et al. (2022) indeed demonstrate that refugees who resettled in Copenhagen experienced 35% faster wage growth for each additional year of experience between 1986 and 1998, surpassing the wage growth of refugees settled in other regions of Denmark. Refugees may therefore see the ONA as an additional tool to help them relocate and to search for better jobs in urban areas. In other words, obtaining the ONA can be a mechanism that mobilises refugees to denser markets resulting in an increase in their employment opportunities. In contrast, since the ONA specifically aims to enhance the labour market participation of refugees, it can contribute to the retention and adaptation of refugees within their initial assignment region. By improving their employability, refugees with the ONA are more likely to find suitable job opportunities, integrate into the local labour market and reduce the need for relocation to other areas in search of employment[29]

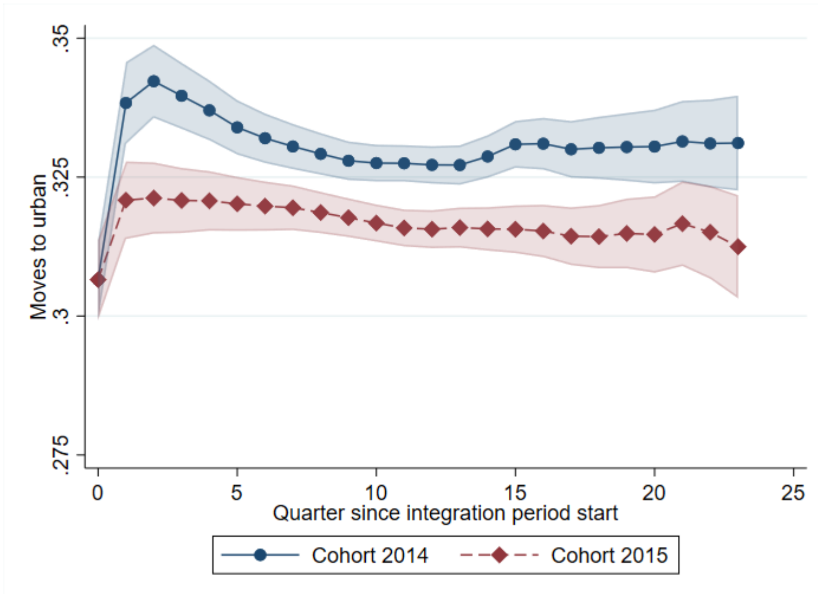
To test these arguments, we construct a variable, *mobility*, that aims to capture the geographic relocation of refugees in the study period. For each refugee, we count the number of location changes over the sample period since obtaining the asylum residence permit in which refugees moved from less urban to more urban areas, using the five categories of density levels that are provided in our data [30] Hence, our dependent variable in these estimations indicates the number of realised location changes of a refugee towards more urban areas from time $t + 1$ for the treatment and control groups.

Our findings, as shown in Figure 2.10, indicate that the ONA is a significant determinant of the geographic mobility of refugees. We find that, for all refugees following an initial period of relocation immediately after obtaining the asylum residence permit (i.e. relocation

[29] Similarly, Foged and Werf (2022) find a negative correlation between language training and geographic mobility, meaning language training reduces the likelihood of moving to more urbanised areas.

[30] According to the Statistics Netherlands, urbanisation categories are defined into five groups: Very strong (2,500 and over/km2); Strong (1500 to 2500/km2); Moderate (1000 to 1500/km2); Weak (500 to 1000/km2); and Non-urban (less than 500/km2).

Figure 2.10: Mobility to urban - Number of moves from rural to urban areas



Note: Based on Equation (2.2) with the number of moves to more urban municipalities being the dependent variable. Estimations include individual, year-quarter-country, and municipal fixed effects as well as individual time-varying characteristics.

from reception centres to social housing), refugees with the ONA are statistically less likely to be mobile towards more urban areas over time by approximately 2 percentage points. This finding is in line with Foged and Werf (2022) who find that linguistic training makes refugess *less* likely to move to urban areas. It is also consistent with our job stability estimations which suggest better trained refugees are more likely to find employment in their local area, change jobs fewer times and stay with each employer for longer. Therefore, relocation to denser labour markets does not correlate with the increased employment probability of those who passed the ONA.

2.6 Robustness Checks

To provide greater confidence in our results, we undertake a series of robustness checks. First, we follow Foged et al. (2022) and apply a complementary Regression Discontinuity (RD) approach.[31] Table B.7 presents the average impact of the ONA over the whole period and confirms the positive effect of the ONA on employment and the lack of any effect on hours worked and hourly wages. The RD estimation shows that completing the ONA component boosts employment probabilities by 4.5 percentage points on average over the whole period, an estimation that falls within the range of period-specific impacts from the baseline specification presented in Table B.5. To compare these findings with our baseline estimates and to see the effect of the ONA over time we plot RD estimates by quarters. Figure B.2, presents these and suggests a significant impact on employment probability around 15 quarters after applying for asylum. By the end of the period, refugees who have completed the ONA have a 10 percentage point higher probability of being employed in comparison with the control group. Both the average and the quarterly effects based on the RD estimation are consistent with our main results.

In another set of checks, we test the robustness of our findings to the matching technique used for PSM. For our baseline specification, we constructed a matched sample on the basis of a balanced panel that excludes refugees who have left the Netherlands during the period of analysis. As a robustness check, we examine whether our baseline results are supported

[31] The specification details and results are presented in Appendix B.6.1.

using different samples: a matched sample based on Coarse Exact Matching (CEM)[32] and an unbalanced panel (i.e. including the leavers). Figure B.3 shows the results, which are again consistent with our baseline findings although the magnitudes from the estimations based on the unbalanced panel (*Panel B*) are slightly lower than those from the baseline estimations.[33].

In a final set of robustness checks, we replicate our baseline estimation for working hours and hourly wages on sub-samples that exclude outlier observations of these variable. More specifically, we drop the highest and lowest 1%, 5% and 10% of observations for these two dependent variables. The results, presented in Figure B.4, confirm our main findings.[34]

2.7 Conclusion

This paper examines the impact of the introduction of the Orientation on the Dutch Labour Market (ONA) component of the Dutch civic integration exam on the economic performance of refugees in the Netherlands. Utilising a unique policy setting with a clear cut-off date for the implementation of this policy and quasi-random allocation of refugees, we compare the

[32] As CEM works by temporarily coarsening each variable into meaningful groups, performing an exact match on these coarsened data, and keeping the original values of the matched data (Blackwell et al., 2009), this technique results in improved matching but at the expense of a substantially reduced sample.

[33] The estimations based on the unbalanced sample indicate that by the end of the analysis period, the treated group has a 4.03 percentage point higher probability of being employed in comparison with the control group.

[34] In an additional set of robustness checks we focus on the precise timing at which each individual passed the exam and utilise a Differences-in-Differences approach that analyses the difference in labour market outcomes before and after passing the integration exam between the treated and the controls. We again find a positive and significant impact of completing the ONA on labour market participation but no impact on hours of worked or hourly wages. These results are available from the authors upon request.

labour market trajectories of individuals who completed the ONA component with those who passed the older civic integration exam without the ONA.

Using linked employer-employee administrative data on the universe of refugees between 2014 and 2021, our main findings indicate that while the ONA increased the probability of employment for refugees, it did not lead to an improvement in their job quality. The divergence in employment probability grew throughout our period of analysis reaching 6.4 -11.0 percentage points after eight years but we do not find any evidence to suggest that the ONA influenced refugees' working hours or hourly wages. We also find that the benefits of the ONA were predominantly observed among male refugees and those from countries other than Syria. These findings are robust to taking unobserved firm effects and firm heterogeneity into account. We find some evidence to suggest that refugees with the ONA are more likely to work for larger, lower labour share firms, although this effect is statistically insignificant between the two cohorts, reinforcing our baseline findings. To put these results in context, Foged et al. (2022) find that refugees' English language training boosted Danish labour market participation by 5.6 percentage points. This comparison therefore indicates that ONA employment training increases employment participation by an effect potentially larger than language training alone, yet it is still not sufficient to secure high-quality jobs without additional skills training.

We further show that the impact of the ONA on employment probability appears to stem from a genuine ‘learning effect’, as those who underwent the ONA demonstrated greater employment stability in two dimensions: i) Firstly, they experienced fewer job/employer changes and a longer average duration of employment; ii) Second, they were less likely to move away from the initial assignment region, both indicating an enhanced matching of workers to firms. Finally, we find only limited evidence that refugees with the ONA are more likely to work for high-wage firms which, we argue, suggests that the ONA does not produce a signalling effect.

While our analysis indicates that the ONA component enhances refugees’ labour market knowledge and their ability to seek and apply for jobs, it also underscores a significant limitation—it fails to equip them with the skills necessary to secure higher-quality employment opportunities and mobility towards more urban areas. Consequently, two potential unintended consequences arise. Firstly, the ONA may lead to a situation where many refugees become overqualified for the positions they hold. Secondly, it could potentially intensify competition for low-skilled jobs, perhaps crowding out native workers, particularly in urban areas where the majority of refugees are located. Moreover, our results indicate that the ONA does not reduce the labour market disadvantages experienced by female and Syrian refugees, as documented in other studies.[35] This necessitates targeted policy interventions that specifically address the mental health issues prevalent among Syrian refugees and the lack of work experience among female refugees.

[35] Earlier studies documented greater prevalence of mental health issues amongst Syrian refugees and a lack of work experience amongst the female refugees. (Helgesson et al., 2019; Bevelander, 2020; Fadhlia et al., 2022)

Based on these findings, this paper offers three key conclusions for policymakers. Firstly, civic integration requirements, as implemented in the Netherlands, do contribute to economic integration among refugees. Secondly, a dedicated requirement focusing on the demonstration of the active use of host-country specific knowledge, as required by the ONA, can enhance refugees' ability to participate in labour markets. However, thirdly, while such a training programme may improve employment prospects, it does not seem to address the issue of low job quality for refugees. This suggests that other factors, such as skills mismatch or labour market discrimination, may continue to limit the types of jobs accessible to refugees. In light of these findings, policymakers should consider addressing these specific challenges to create more inclusive labour markets, while also exploring strategies to bridge the gap between employment rates and job quality for all refugees.

Chapter 3

Firm Sorting and the Economic Assimilation of Refugees

Abstract

The literature on immigrant-native pay gaps points out that such gaps can be explained not only by individual (un)observed characteristics, but also by firms via two complementary channels: firm sorting; and pay-setting. Firm sorting focuses on differences in the fractions of natives and immigrants employed in a firm with specific pay premiums, while the second channel emphasises the differences in wage-setting between natives and immigrants within firms. In this context, we investigate whether and to what extent firm sorting effects influence the native-immigrant wage gap, using linked employer-employee data for 2014-2022 in the Netherlands. Focusing on native-*refugee* gaps and using a two-way effects model (the AKM model), we show significant pay differentials between refugees and natives, with refugee wages approximately half of those of natives. We also find a substantial mismatch between ability and firm-specific pay premiums for this group. Our analysis also suggests that individual characteristics are the main drivers of pay gaps *within* and *between* groups, though firm sorting accounts for 14% of the native-refugee pay gap. The pay-setting effect is close to zero. The adverse impact of firm sorting may stem from refugees having limited access to higher-paying firms upon entering the labour market and may result in a mismatch between refugees' abilities and wages.

Keywords: Refugee; economic assimilation; pay disparity; firm effect

JEL classification codes: J15, J31, J61

3.1 Introduction

The difficulties experienced by refugees in the labour market have been well-documented and stem from factors such as disrupted education or working tenures, and poor physical or mental health perhaps as a result of trauma and discrimination. Refugees are typically less likely to secure employment and more likely to experience persistent long spells of unemployment. When refugees are employed, their jobs tend to be low-paid and they are more likely to be over-qualified for the jobs they hold (Bevelander, 2020; Brell et al., 2020). Fasani et al. (2022), for example, estimate that the unemployment rate among refugees may be 22% higher than that of other migrants with similar characteristics, and these gaps persist even 10 years after migration.

Government policies towards refugees can compound these difficulties by limiting access to suitable labour markets and co-ethnic networks. Indeed, policies such as random regional allocation, lengthy asylum applications, employment bans during the application process and restrictions to only selected industries or occupations have been shown to amplify the labour market participation gap between refugees and natives or other immigrants (Edin et al., 2004; Hainmueller et al., 2016; Fasani et al., 2021; Ahrens et al., 2023). However, appropriate assistance programmes, such as local language training and active labour market policies, can help to improve refugees' language skills and to find better-matched jobs thereby improving their labour market performance (Auer, 2018; Foged et al., 2022).

Although the under-performance of refugees in labour markets is well documented, little is known about how the types of jobs refugees undertake contribute to their economic integration. To the best of our knowledge, Baum et al. (2020) is the only study to examine the occupational sorting of refugee workers. Using Swedish data and focusing on established refugees, Baum et al. (2020) show that occupational sorting is an important driver of the wage gap between refugees and natives. Refugees are more likely to be sorted into low-skilled occupations than comparable native workers. In contrast, occupational differences play almost no role in explaining the gender pay gap once firm heterogeneity is taken into account (Cardoso et al., 2016). These findings suggest the impacts of refugee firm heterogeneity may explain why refugees are disadvantaged in the labour markets. These results echo those of Chapter 2, which finds that although labour market training programmes improve the labour market participation of refugees, they do not lead to longer hours worked or higher wages (Yumoto et al., 2023b). These findings suggest that refugees are stuck in low-paid jobs and are likely to experience a persistent wage gap with natives and with other immigrant groups.

The literature examining the pay gaps between immigrants and natives, and gender groups, points out that such gaps can be explained not only by individual (un)observed characteristics but also by the characteristics of firms where individuals are employed (see, for instance, Aslund et al. (2021) and Dostie et al. (2021) for the native-immigrant pay gap, and Card et al. (2016, 2018); Li et al. (2023) for gender pay gaps). These studies consistently point out that individual characteristics such as gender, age and educational attainment are important determinants of wages, however firm-specific effects are also substantial. For instance, Cardoso et al. (2016) show that firm effects explain approximately 20% of the gender pay gap and claim that such effects include firms' wage policies and the tastes of employers, customers, or co-workers, including the possibility of discrimination. Manning and Swaffield (2008) show that occupations play a limited role in explaining the expansion of gender pay gaps in the UK from the stage of entering into the labour market until a decade later. Damas de Matos (2017) finds approximately 30% of the native-immigrant wage catch-up is linked to worker mobility into firms with a greater wage premium. This may be due to immigrants having little knowledge of the labour market which limits their ability to identify firms with higher pay, while simultaneously higher-paying firms may be reluctant to hire immigrants without local labour market experience.

A number of studies find that firm heterogeneity affects the pay gap via two complementary channels: firm sorting; and pay-setting (e.g. Card et al. (2016); Aslund et al. (2021); Dostie et al. (2021)). Firm sorting focuses on differences in the fractions of natives and immigrants employed at firms with firm-specific pay premiums. Firm sorting widens the native-immigrant pay gap when immigrants are underrepresented in firms that offer a higher pay premium. In other words, when natives are concentrated in firms with higher pay and immigrants are more likely to be employed at low paying firms the pay gap is likely to be widened via the firm sorting channel. If natives and immigrants were equally dispersed in firms in terms of firm-specific pay premiums, the sorting channel would not contribute to the pay gap. The second channel represents the differences in wage-setting between natives and immigrants within firms and would expand the pay gap when immigrants earn less than natives at a given firm. Firm sorting reflects gaps *between* firms while pay-setting reflects gaps *within* firms.

Studies of Portugal and Italy show that firm effects explain approximately 21% and 34% of the gender pay gap, respectively, mainly through firm sorting effects (Card et al., 2016; Cardoso et al., 2016). Furthermore, Cardoso et al. (2016) point out that while the gender wage gap has shrunk over time, the gap in firm-specific premiums remained unchanged, which increases the contribution of firm effects to the overall pay gap. Focusing on the immigrant-native gap, Arellano-Bover and San (2020) show that the assimilation of immigrants stems from climbing the job ladder into firms with a greater pay premium as time in the host country increases, which contributes to a reduction in the wage gap. Dostie et al. (2021) show that firm sorting effects explain 20% of the immigrant-native pay gap while the impact of pay-setting is essentially zero in Canada.

This strand of the literature has, however, paid little attention to refugees despite the particular labour market challenges they face. We aim to address this by undertaking a comprehensive analysis of firm sorting and pay-setting effects on the wage gap between natives, second-generation immigrants, first-generation immigrants and refugees. Using Dutch employer-employee data from 2014 to 2022 we apply a two-way effect model originally suggested by

Abowd et al. (1999) (hereafter, the AKM model) to identify the extent to which firm effects can explain wage disparities *within* groups, and a Blinder-Oxaca style decomposition model to distinguish the contribution of firm-sorting and pay-setting effects to the wage gaps *between* groups (Card et al., 2016; Dostie et al., 2021).

The Netherlands is representative of the wider EU in terms of the level of refugees per capita and the composition of recently arrived cohorts and therefore offers an ideal setting to analyse the integration of refugees in the European context as shown in Chapter 1 (Yumoto et al., 2023a). Nevertheless, some aspects of the Dutch integration process differ from those elsewhere in the EU. For instance, asylum-seeking applications are processed at a faster pace in the Netherlands than in other European countries. On average, asylum seekers in the Netherlands can expect an asylum decision after 4.5 months compared to more than 9 months in Germany (Bertoli et al., 2020) or in the United Kingdom where two-thirds of asylum applicants waited for more than 6 months. Such a relatively short application processing time may help to alleviate adverse impacts, skill atrophy and physiological discouragement, as suggested by Hainmueller et al. (2016). While asylum seekers are not allowed to work within the first six months of arrival, thereafter they are allowed to work without restrictions in most sectors with an Asylum Residence Permit.[1] Moreover, refugees in the Netherlands are randomly allocated to social housing in municipalities.[2] This regional dispersal policy may limit

[1] Some public positions, such as mayor, police officer, soldier or judge, are only available for Dutch citizens (https://ind.nl/en/advantages-and-disadvantages-of-becoming-a-dutch-citizen).

[2] Refugees are free to move out to different locations from initially allocated municipalities only if they are economically independent to pay rents and are prepared to leave social housing. Yet, due to their economic challenges, more than 80% stay in the initially assigned municipalities in our sample.

access to appropriate local labour markets, and thus, hamper refugees' labour market performance (Edin et al., 2004; Ahrens et al., 2023).

Our analysis confirms that refugees experience significantly lower wages with a smaller standard deviation compared to natives and other immigrants. While native workers earn EUR 20.25 per hour with a standard deviation of EUR 10.29, refugees are paid EUR 10.86 with a standard deviation of only EUR 3.63. We observe that highly paid refugee workers (i.e. high ability refugees) obtain lower firm-specific pay premiums, indicating that refugees suffer from a skill-wage mismatch. Our analysis suggests that wage disparities are primarily driven by individual characteristics rather than firm pay premiums with the latter being mainly driven by sorting effects rather than pay-setting effects. The individual observed and unobserved characteristics account for 64-71% of the total wage disparities within groups, and explain approximately 80% of the wage gaps between groups. More specifically, pay-setting effects make a minimal contribution to the native-refugee wage gap, which indicates that once refugees secure a job, they are paid equally to their native co-workers. The limited access of refugees to higher-paying firms is also confirmed by the fact that refugees' native co-workers have lower average wages (EUR 17) compared to natives as a whole (EUR 20), and we observe a similar pattern for the wages of co-workers of first-generation immigrants.

The remainder of this paper is structured as follows; Section 2 presents the data and summary statistics and Section 3 presents our empirical methodology. The empirical findings are presented and discussed in Section 4 and Section 5 concludes.

3.2 Data and sample

3.2.1 Data source and sample selection

We base our analysis on Dutch administrative datasets provided by Statistics Netherlands (CBS). Our main dataset is employer-employee data, which consists of *SPOLISBUS* and *Productiestatistiek*. *SPOLISBUS* provides monthly data on employers, earnings, and hours worked for the population of employed individuals in the Netherlands, whereas *Productiestatistiek* is annual firm census data containing financial information such as value-added and sales at the firm level.[3] The two datasets are merged with a unique firm identifier. Furthermore, we complement the employer-employee data with personal demographic information such as age, gender and country of birth with *BRP* (Personal Record Database), and administrative information on the universe of refugees given asylum since 2014 (i.e. *Asielcohort*), to distinguish refugees from other types of immigrants. We combine employee level variables with employer level data using the unique social security number of each individual. Our sample is restricted to employed individuals of working age. We distinguish between

[3] The sample of *Productiestatistiek* covers approximately 10% of the total Dutch business population until 2021 (all of the financial data is used as the *average* values for the whole period as shown in Equation 3.3). For companies with fewer than 10 employees, the information is obtained from records of the tax registrations or by means of surveys on a sample basis. Companies with 10 to 50 employed persons are randomly approached with a questionnaire. Companies with 50 or more employed persons all receive a survey form (https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksomschrijvingen/productiestatistiek).

four groups based on their immigration status; natives (*autochtoon*); second-generation immigrants; first-generation immigrants and refugees. Since refugees who arrived before 2014 cannot be distinguished from first-generation immigrants due to data availability, we limit our analysis to the period between 2014 and 2022.[4]

For our main specifications and to minimise the impact of outliers, we winsorise at the 1st and 99th percentile of the hourly wage distributions of the whole sample in each year. In addition, we drop observations related to individuals whose wage (in log terms), in any year, is three standard deviations or more above the sample mean (Aslund et al., 2021). We adopt a similar approach and winsorise firm-level data at the 5th and 95th percentile of value-added and sales. This results in a total sample of 55,722,555 observations for natives, 6,987,642 observations for second-generation immigrants, 2,481,491 observations for first-generation immigrants and 217,875 observations for refugees.

In order to distinguish firm fixed effects from individual fixed effects, we exploit a 'connected set of firms', which refers to firms linked by worker mobility, and workers at the connected firms, as suggested by Abowd et al. (2002). Firms are connected if the firm employs at least one worker observed working for another company at a different point in time. In our following analysis, we limit attention to workers and firms in the largest connected set

[4] We identify the immigrant generation mainly based on the date of arrival in the Netherlands, complementing the parents' nativity (i.e. foreign-born), which is identified through *BRP*. We drop first-generation immigrants who arrived before 2014 and keep second-generation immigrants, who have at least one foreign-born parent, born in the Netherlands.

for each group to estimate wage gaps *within* groups. When comparing wage gaps *between* groups and estimation gaps *within* groups, we further narrow down our sample to workers employed at firms in the connected sets for both groups (i.e., dual-connected set). For instance, to specify firms' contributions to the wage gap *between* natives and refugees, we exploit the sample firms which are connected by native workers and also linked by refugees, hence, where native and refugee workers are employed in the dual-connected set.

One limitation of our data is a lack of information on the educational attainments of employees. Dostie et al. (2021) show that controlling for education and age reduces the observed pay gap but leave the changes in the pay gap over time largely unaffected. Their study suggests that an analysis of wage gaps without adjusting for education remains informative. Moreover, given that participation in formal educational activities is low among adults in the Netherlands (OECD, 2019a), education is relatively constant for most workers, thus, we can expect that the unobserved impact of education will be captured by individual-level fixed effects. 1.7% of the individuals in our sample with educational information obtained different levels of education for the period. As a robustness check, we add education information, using an additional dataset called *HOOGSTEPL* with data available only until 2021, and confirm that our main results are consistent regardless of whether we control for educational attainment or not.[5]

[5] To fully confirm the assumption for our model (AKM model), we need to keep observations in 2022, thus, dropping educational information.

3.2.2 Descriptive Analysis

We start with an overview of the wage gap between natives and different groups of immigrants. Figure 3.1 reports the result of a simple OLS estimation[6] and shows the ratio of the hourly wage between natives and each of the immigrant groups over our time period. Panel A compares natives and second-generation immigrants and indicates that once we control for individual-level characteristics such as age and gender, industry effects and firm effects, the wage gap significantly narrows from EUR 1.16 (i.e. 0.152 log points) without any controls to EUR 1.04 (i.e. 0.0396 log points) in 2014. Panel B presents the gap between natives and first-generation immigrants and shows that even after adding a full set of controls, first-generation immigrants earn EUR 1.26 (i.e. 0.23 log points) less than natives in 2014. However, in Panel B we also show a declining raw pay gap over time that could potentially be due to the accumulation of host country-specific human capital, such as an improvement in Dutch language skills or obtaining degrees and qualifications in the Netherlands.

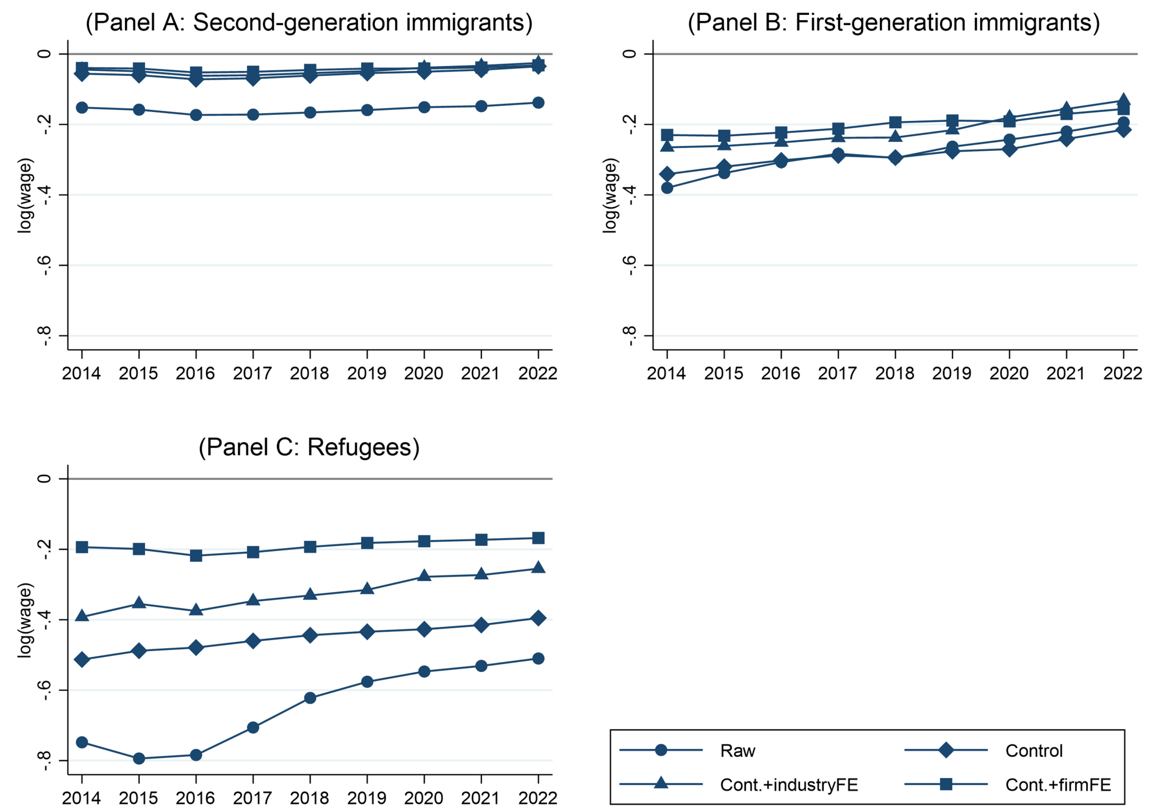
Panel C compares natives and refugees and indicates a considerable wage-gap between the two groups. When we incorporate additional controls into the analysis we observe a smaller pay gap, with the specification that includes both individual controls and firm fixed effects yielding the smallest gap. This specification substantially reduces the pay gap (i.e., from 0.78

[6] The simple OLS results stem from the following equation

$$\ln{(w_{it})} = c_t + \theta_t imm_i + \beta_t X_{it} + \upsilon$$

Where, θ_t are log hourly wages for worker i in year t at the primary employer, X_{it} refer to controls for age, age squared and gender as well as industry dummies or firm fixed effects and υ is the residual.

Figure 3.1: Trend of wage disparities with natives



Note: θ_t from the following naive regressions: $\ln(w_i t) = c_t + \theta_t imm_i + \beta_t X_{it} + \upsilon$ Where θ_t are log hourly wages for worker i in year t at the primary employer. Included in X_{it} are controls for age, age squared and gender as well as industry dummies or firm fixed effects. Raw shows θ without any control, and Controls include age, age squared and gender. These estimations are based on the whole sample.

to 0.19 log points in 2014 and from 0.51 to 0.16 log points in 2021, respectively). These findings suggest that firm-specific premiums play a significant role in explaining the pay gap between native and refugee workers in addition to individual characteristics. Furthermore, for the other immigrant groups in Panels A and B, the regressions with the individual controls and firm effects show the smallest gaps in most years, indicating these factors have a substantial role in the pay gap even though the magnitudes vary.

Figure C.1 shows the composition of nativity groups in each hourly wage decile. Again, it suggests natives outperform all types of immigrants. Natives are concentrated in higher wage deciles, more specifically, natives form over 90% of 6th deciles and above. The lowest decile is the one second-generation immigrants are most concentrated in and, as the wage increases, their composition shrinks to about 8% in the 10th decile. First-generation immigrants and refugees are concentrated in the second lowest decile, and have almost no representation in the 6th wage deciles and over.

Figure C.2 highlights the wage growth and gaps of first-generation immigrants and refugees as they spend time in the Netherlands. As Albert et al. (2021) report, recently arrived immigrants experience considerable wage gaps with natives, although they also experience rapid wage growth and converging wage gaps for the first few years in the host countries. Panels A and B present the wage gaps relative to native wages. While Panel A shows the gaps in hourly wages at the individual level, Panel B presents the gaps in the average hourly wages of the firms in which individuals work, which indicates whether the individual is employed at a higher wage paying firm or not. Both panels indicate that the wage gaps have reduced for first-generation immigrants and refugees as they climbed the job ladder in terms of the average wage of firms as time in the Netherlands passes. Panels C and D illustrate wage growth relative to wages in the first year in the Netherlands, suggesting that both groups enjoyed rapid rises over the first 6-7 years.

Table 3.1 shows the summary statistics by groups for the overall sample and for the sub-samples of connected sets. It indicates that all types of immigrants are on average younger than natives by approximately 10 years though refugees are substantially younger (i.e. 28 years old). Moreover, males are disproportionately over-represented among refugees (over 80%). In terms of hourly wage, natives have on average higher earnings in comparison with all groups of immigrants. Although the wage gap is the narrowest between natives and second-generation immigrants it is striking that such a gap is present given that second-generation immigrants are expected to be fluent in the Dutch language and to have sufficient levels of local human capital. Table 1 shows a substantial wage gap between refugees and natives; on average a refugee earns approximately half the hourly wage of a native. Limiting the sample to the connected sets of firms results in a significant reduction in the sample size (between 1.5 to 8.1 per cent of each group), but these sub-samples reflect the overall sample for each group in terms of their characteristics.

Table 3.1: Summary statistics

	Overall Analysis Sample				Connected Sets of Worker/Firms			
	Native (1)	2nd-gen (2)	1st-gen (3)	Refugee (4)	Native (5)	2nd-gen (6)	1st-gen (7)	Refugee (8)
Individual level								
Age	40.01	33.71	31.25	28.50	39.90	33.47	31.18	28.26
(std.dev.)	14.04	13.26	8.75	9.32	14.04	13.16	8.70	9.23
Female	48.24%	49.23%	44.55%	19.86%	48.50%	49.50%	44.65%	19.66%
Hourly wages	20.25	17.64	15.63	10.86	20.17	17.55	15.60	10.84
(std.dev.)	10.29	9.86	8.79	3.63	10.22	9.76	8.74	3.61
Number of								
person-year	55,722,555	6,987,642	2,481,491	217,875	55,058,240	6,842,168	2,429,168	207,035
persons	7,946,874	1,117,095	726,811	68,577	7,828,891	1,088,326	704,689	63,013
Firm level								
Firm size	16.58	46.09	100.03	206.68	19.59	51.89	113.79	231.38
(std.dev.)	261.23	462.90	804.29	1,368.33	285.35	493.96	865.53	1,469.82
Immigrant share	10.23%	41.81%	55.11%	54.50%	11.58%	39.82%	54.23%	55.14%
(std.dev.)	18.68%	33.09%	32.82%	34.92%	19.32%	32.04%	32.24%	34.62%
Number of firms	692,962	260,115	96,535	30,121	595,706	234,058	79,594	25,343

Note: Hourly wages are deflated with the base of the 2015 price rate. Firm size is measured with full-time equivalent employee units. The summary statistics in dual-connected set are shown in Appendix.

Table 3.2: Industry distribution by groups

	All (1)	Native (2)	2nd-gen (3)	1st-gen (4)	Refugees (5)
Health, mental and social care	17.47%	18.54%	14.50%	2.84%	4.33%
Government, education and science	6.92%	7.24%	5.32%	4.78%	0.65%
Business services	5.84%	5.61%	7.53%	6.53%	2.98%
Hospitality industry	4.86%	4.33%	7.25%	8.58%	20.95%
Lending companies	4.03%	3.04%	6.21%	18.83%	16.15%

Note: The table shows the distribution of main industries by groups within the overall analysis sample. The industry categories are defined by the Tax and Customs Administration.

Table C.1 in the Appendix provides summary statistics for the dual-connected sets of firms. Focusing on the dual-connected set reduces the sample size even further. The number of observations for natives in the dual-connected set of natives and refugees is 44.0% of the total sample. Moreover, the hourly wages of natives in the dual-connected are significantly lower in comparison with the overall sample. While the average hourly wage of native workers in the overall sample is EUR 20.25, the corresponding value in the dual-connected set declines to EUR 16.75. The hourly wage of refugees remains at the same level, approximately EUR 10.8, in the overall sample and the dual-connected set. This suggests that native workers at firms where refugees are employed earn significantly less than the average native population. In other words, refugees do not seem to have access to firms with higher firm-specific pay premiums, unlike other workers. This descriptive analysis implies that firm sorting effects contribute to a widening of the native-refugee pay disparities.

3.3 Estimation strategy

To identify firm effects on wage disparities between natives and immigrants, we proceed in two steps. First, we apply a Two-Way effect model, known as the AKM model, originally suggested by Abowd et al. (1999, 2002) and extended for decomposition by Card et al. (2016) and Dostie et al. (2021). While the AKM model identifies the contribution of firm effects to the wage gaps *within* groups, the decomposition results indicate the extent to which firm effects contribute to wage gaps *between* groups. These firm effects from the decomposition model can be decomposed into our two channels (i.e., sorting and pay-setting effects).

3.3.1 AKM models (Two-way effect models)

Our approach adopts a simple Two-way model that allows us to determine the effect of firm-specific pay premiums on the observed wages of each group. We assume that we observe point-in-time wages for workers (i.e., $i \in \{1, ..., N\}$) in multiple periods $t \in \{2014, ..., 2022\}$. We denote the group of workers i by $g(i)$ which takes the value N for native, S for second-generation immigrant, F for first-generation immigrant, or R for refugee, and the identity of worker i's employer in a given year by $J(i,t)$. The baseline model of the real hourly wage (in natural logarithms) (i.e., $\ln(w_{it})$)[7] earned by individual i from group $g \in \{N, S, F, R\}$ in year t is given by:

$$\ln(w_{git}) = \alpha_{gi} + \psi_{J(git)} + X_{git}\beta + \varepsilon_{git} \tag{3.1}$$

[7] The hourly wages are deflated to the 2015 price level.

Where α_{gi} captures the person effects of individual i from group g, $\psi_{J(it)}$ refers to the firm effects of firm J where individual i works in year t, X is a vector of time-varying controls such as age, and includes year fixed effects, and ε refers to the residuals. To avoid perfect collinearity between year and age, we follow Card et al. (2018) and use quadratic and cubic terms, normalising at the age of 35 (i.e. $(a_{it} - 35)^2 + (a_{it} - 35)^3$). Age 35 is relatively close to the average age in the different groups and at this age, the age profiles of wages, which refers to wage trajectories over ages, tend to be relatively flat across groups (see Figure C.3 in the Appendix) [8] For this specification, we exploit the connected set of firms, which allows us to distinguish firm effects from person effects.

To specify the contribution of firm effects to the whole wage gap within the groups, the variance of the log wage from Equation 3.1 can be decomposed in the following way:

$$\begin{aligned} \mathrm{var}(\ln(w_{git})) &= \mathrm{var}(\alpha_{gi}) + \mathrm{var}(\psi_{J(git)}) + \mathrm{var}(X_{git}\beta) + \mathrm{var}(\varepsilon_{git}) \\ &+ 2\,\mathrm{cov}(\alpha_{gi}\,\psi_{J(git)}) + 2\,\mathrm{cov}(\alpha_{gi}\,X_{git}\beta) + 2\,\mathrm{cov}(\psi_{J(git)}\,X_{git}\beta) \end{aligned} \tag{3.2}$$

This specification identifies the contribution of each term to the wage variance within a group, as well as the covariance between person and firm effects which indicates whether high ability individuals are more likely to be employed at higher paying firms. We are particularly

[8] Although Card et al. (2018) set 40 to be the age profile as the baseline, they also confirm that firm effects are non-sensitive to the chosen age. They estimate the AKM model with different ages of 30, 50 and 0 (i.e. $(a_{it} - 30)^2 + (a_{it} - 30)^3$, $(a_{it} - 50)^2 + (a_{it} - 50)^3$ and $(a_{it})^2 + (a_{it})^3$), and the contribution of firm effects remain unchanged. We also replicate the baseline AKM model with different ages (Table C.5).

interested in $\psi_{J(git)}$ to examine whether refugees' wages are associated with firms more than the other groups (i.e., $\psi^R > \psi^N$, ψ^S, ψ^F).

3.3.2 Exogenous mobility assumption

Card et al. (2013, 2016, 2018) and Dostie et al. (2021) point out that the AKM model based on OLS yields unbiased estimators only when the assumption of exogenous mobility among firms is met. Exogenous mobility means that worker mobility is uncorrelated with the time-varying residual components of wages.

The exogenous mobility assumption can be violated through three channels. The first channel is related to job-match components, which refer to idiosyncratic combinations of jobs and workers. If the job-match components are dominant determinants of workers' mobility across firms, all moves will lead to wage gains as in the dynamic matching model of Eeckhout and Kircher (2011). In the absence of job-match effects, the mean bias associated with joiners and leavers would cancel out at the aggregate level, as would occur when firms' employment is in a stable state. The second channel is when workers' mobility is driven by firm-wide shocks. For instance, workers tend to leave firms having negative shocks for firms experiencing positive shocks. In this case, we would expect to observe a systematic fall in the wages of leavers just before the move and unusual wage jumps for joiners. The third channel is observed when the direction of workers' mobility across firms is correlated with transitory wage shocks. Exogenous mobility is violated through this channel when people moving to higher-wage firms have different trends prior to moving than those who move to lower-wage firms.[9]

To check the validity of the assumption of exogeneous mobility, we undertake simple event studies as suggested by Card et al. (2013, 2016, 2018); Dostie et al. (2021). We select job-changers (i.e. movers) with at least two consecutive years in both origin and destination firms and classify origin and destination firms in wage quartiles. We classify origin firms on the basis of the average co-worker wage (i.e the average hourly wage at the firm excluding the wage of the mover) in the last year at this firm, and the destination firms on the basis of the mean co-worker wage in the first year when the movers join this firm. We then assign movers to the 16 categories defined by the combination of the quartiles of origin and destination workplace.[10]

We can confirm exogenous mobility if we observe approximate symmetry of the wage trajectories among movers and flat profiles in the wages before and after the job change. Approximate symmetry of the wage trends refers to the mean changes in unobserved determinants for joiners of firm j and the corresponding change for leavers of this firm are

[9] According to Card et al. (2016), in this violation, workers who are performing well can be more likely to transit to firms with higher premiums, whilst workers whose performances are poor would tend to downgrade firms in terms of pays. In the absence of this violation, the two types of movers would have similar and flat trends before the transition.

[10] While Card et al. (2016) group all workers of both genders into 16 origin-destination quartiles, we classify workers within groups. This is because we cannot observe first-generation immigrants and refugees that are employed at the top-quartile firms of all workers, but have the sample of the top-quartile firms within each group. In addition, as the AKM model is estimated *within* group, all we need to confirm is the exogenous mobility within groups.

roughly comparable in magnitude since the decision to leave one firm is a decision to join another (Card et al., 2016). For example, the value of gains to movers from the lowest quartile to the highest quartile should be approximately equal to the magnitude of loss for movers from the highest to the lowest quartile. If mobility is endogenous it will be linked to the match component of wages. If this is the case, movers will tend to experience positive wage gains regardless of the direction of their move which violates the symmetry prediction (Card et al., 2018). If we observe flat wage profiles before and after mobility, where wages are all relatively stable except in the transition years, we can confirm the absence of the violations associated with firm-wide shocks and transitory wage shocks. Workers who are about to experience a major wage loss by moving to a firm in a lower co-worker pay group show no obvious wage trends beforehand. Similarly, workers who are about to experience a wage rise by moving to a firm in a higher pay group also show no such evidence. Flat wage profiles indicate the absence of sudden wage drops for leavers just before their exit and unusual wage growth for recent joiners, and an absence of transitory wage shocks (Card et al., 2016).[11]

3.3.3 Normalising firm effects

Card et al. (2016, 2018) and Dostie et al. (2021) consider that low-productive, or 'zero-surplus', firms do not pay firm-pay premiums to their workers. In other words, firms with productivity below a certain threshold pay no firm-pay premiums, whereas more productive

[11] In the presence of transitory wage shocks, systematic mobility implies that people moving to higher-wage firms will have different trends prior to moving than those who move to lower-wage firms

firms pay positive firm premiums.

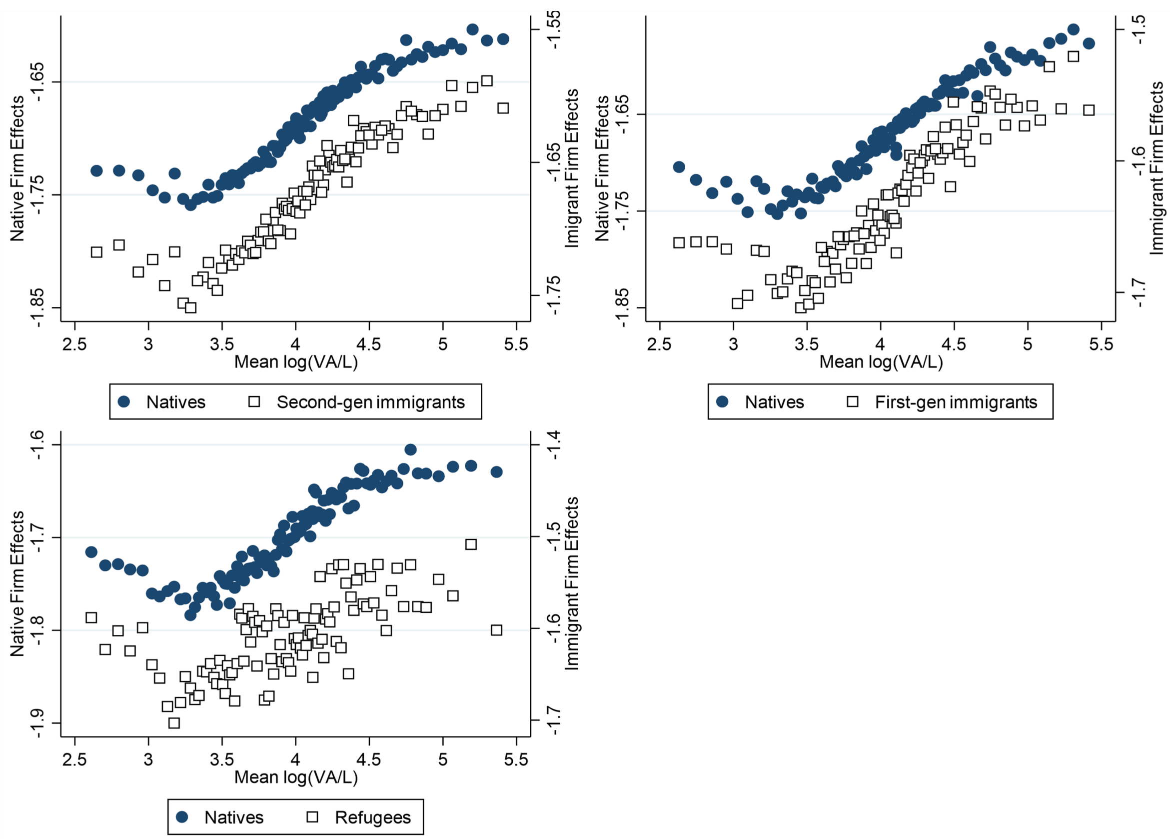
To better understand the relationship between firm productivity and firm-specific premiums, we group firms into percentile bins of value added per worker and plot the average firm effects by group in each bin against the mean log value added per worker of the corresponding bin. Figure 3.2 presents the results and shows that, for all groups, low-productivity firms, below a threshold between 3.5 and 4.0 log-point of value-added per worker, pay similar firm-specific premiums whereas firms with productivity levels above this threshold tend to pay firm-specific premiums that increase with their level of productivity in a linear way.

To specify the productivity threshold in a statistical way, we apply the following specification (Equation 3.3) for each group using firm-level data (Card et al., 2016; Dostie et al., 2021):

$$\begin{aligned} \widehat{\psi}_j^N &= C_N + \delta'_N \max\{0, (\bar{S}_j - \tau) + e_j^N \\ \widehat{\psi}_j^M &= C_M + \delta'_M \max\{0, (\bar{S}_j - \tau) + e_j^M \end{aligned} \tag{3.3}$$

Where $\bar{S}_j$ represents the average log of gross value added per worker over the time period for a given firm and τ is the productivity threshold for a specific group. Using firm-level data for the dual-connected set with financial data, we estimate Equation 3.3 with different values of τ and select a value that maximises the goodness of fit of the model (i.e. R^2 and the mean

Figure 3.2: Firm effects versus log value added per worker



Note: Points shown represent mean estimated firm-specific wage premiums from the AKM model for natives and immigrants in percentile bins of mean log value added per worker.

squared error). Although Equation 3.3 is estimated for each group separately, the value of τ should be identical in the dual-connected set (Dostie et al., 2021). Given the estimated $\hat{\tau}$, we normalise the estimated firm effects for both groups to have an average of zero across all firms with $\bar{S}_j < \hat{\tau}$. [12]

[12] This is essentially the same as subtracting the estimated values of the constants (C_N and C_M) in equation 3.3 for each group (Card et al., 2016).

3.3.4 Decomposition model

The AKM model is useful to explain pay gaps within groups but is not informative about pay gaps *between* groups. To further explore the pay gap *between* groups, we also compare natives and each of the three types of immigrants, setting natives as the reference group.[13] We apply a Blinder-Oaxaca style method to decompose the gaps in firm-specific pay premiums between the groups into a combination of sorting and pay-setting effects. Based on Equation (3.1), let D_{git} indicate whether individual i from a group g, either natives (N) or immigrants ($M \in \{S, F, R\}$), is employed at time t and π_{gjt} be the fraction of each of the two groups at firm J. The mean log wages of natives and immigrants observed in year t can be decomposed as follows:

$$\begin{aligned} E[\ln(w_{Nit})] &= E[\alpha_{Ni}|D_{Nit}=1] + \bar{X}_{Nt}\beta_N + \sum_j \psi_j^N \pi_{Nit} \\ E[\ln(w_{Mit})] &= E[\alpha_{Mi}|D_{Mit}=1] + \bar{X}_{Mt}\beta_N + \sum_j \psi_j^M \pi_{Mit} \end{aligned} \tag{3.4}$$

In order to compare the firm effects between natives and immigrants, we limit our sample to firms in the dual-connected set i.e. firms that are in the connected sets for both native and immigrant groups. For example, when comparing natives and refugees, we use the dual-connected set of natives and refugees. Aslund et al. (2021) and Dostie et al. (2021) suggest comparing mean wages between natives and immigrants by taking a difference in the mean

[13] We separately estimate this specification based on three different dual-connected sets: natives and second-generation immigrants; natives and first-generation immigrants; and natives and refugees.

of log wages between the reference group (i.e., natives) and one of the immigrant groups:

$$\begin{aligned} E[\ln(w_{Nit})] - E[\ln(w_{Mit})] &= E[\alpha_{Ni}|D_{Nit}=1] - E[\alpha_{Mi}|D_{Mit}=1] \\ &+ \bar{X}_{Nt}\beta_N - \bar{X}_{Mt}\beta_M + \sum_j \psi_j^N \pi_{Nit} - \sum_j \psi_j^M \pi_{Mit} \end{aligned} \tag{3.5}$$

The first term on the right side shows the difference in the means of the time-invariant person component, including gender and inborn ability, between the two groups, while the second term represents the gaps in the observed time-varying components such as age. We are interested in the last term: the net contribution of firm-specific wage premiums to the pay gap between natives and immigrants. Following Oaxaca (1973), the net firm effects can be decomposed into the two channels as in Equation (3.6):

$$\begin{aligned} \sum_j \psi_j^N \pi_{Njt} - \sum_j \psi_j^M \pi_{Mjt} &= \sum j \psi_j^N (\pi_{Njt} - \pi_{Mjt}) + \sum j (\psi_j^N - \psi_j^M) \pi_{Mjt} \\ &= \sum j \psi_j^M (\pi_{Njt} - \pi_{Mjt}) + \sum j (\psi_j^N - \psi_j^M) \pi_{Njt} \end{aligned} \tag{3.6}$$

Where the first term, a weighted average of the differences in employment shares of the two groups (i.e. $\sum j\psi_j^N(\pi_{Nit} - \pi_{Mit})$), shows the sorting effects and the last, a weighted average of the differences in pay premiums (i.e. $\sum j(\psi_j^N - \psi_j^M)\pi_{Mit}$), indicates the pay-setting effects. The two effects can be also computed as follows: $\sum j(\psi_j^N - \psi_j^M)\pi_{Mit}$ for sorting effects and $\sum j(\psi_j^N - \psi_j^M)\pi_{Mit}$ for pay-setting effects. Equation (3.6) allows us to identify how much firm sorting and pay setting effects contribute to the native-immigrant pay gap.

3.4 Estimation results

This section presents our findings based on the methodologies described in the previous section. We start by assessing the validity of the AKM model on the basis of the exogenous mobility assumption. We then proceed by presenting the results of the AKM model and of the normalisation of estimated firm effects with productivity thresholds. Finally we present the results from the decomposition model in the last sub-section.

3.4.1 Exogenous mobility

Figure 3.3 presents the results of simple event analysis by groups. We confirm the approximate symmetry of wage changes among movers across groups. Workers who move between jobs with higher- and lower-paid co-workers experience systematic wage gains and losses in an approximately symmetric way. On average, native movers from the bottom quartile firms to the top quartile experience a wage rise of EUR 16.16 (i.e. EUR 10.23 or 2.325 log points at t_{-1} to EUR 26.39 or 3.273 log points at t_0) whilst those who moved from the top to the bottom face a wage reduction of EUR 16.12 (i.e. EUR 27.33 or 3.308 log point to EUR 11.21 or 2.417 log point). Likewise, for the other groups, we can observe approximate symmetries in wage changes among immigrant movers from the top to the bottom quartiles and the bottom to the top.[14] Approximately symmetric wage changes accompanied by firm transitions indicate the presence of significant firm-specific pay premiums, rather than idiosyncratic job-match effects (i.e. mobility patterns are not driven by comparative advantages in wages) (Card et al., 2016).

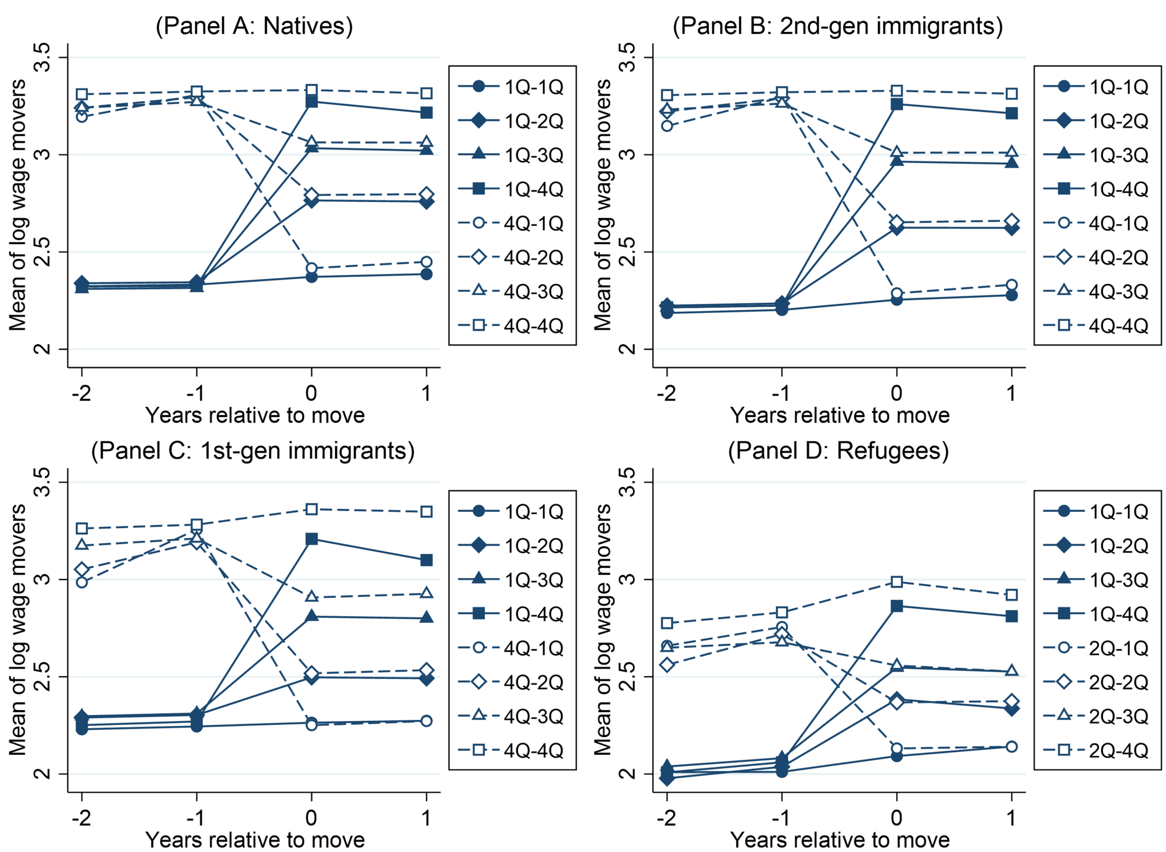
As a whole, we also see relatively stable wage profiles before and after mobility. That is, there is no indication that movers to firms with higher- or lower-paid co-workers have different wage trends *prior* and *subsequent* to the move. Furthermore, wage changes for people who move between firms with similarly paid co-workers show little or no excess wage growth or reductions.[15]

Dostie et al. (2021) suggest that an alternative way to test the exogenous mobility assumption is to compare the goodness of fit of models, the root of MSE and adjusted R^2, between the AKM model (i.e. Two-way effect model) and the job-match effect model which would show whether the job-match effects are dominant in the wage variance. They suggest that the small magnitude of differences between the two models would alleviate concerns over the potential role of job-match effects in accounting for mobility patterns. As a further test of the exogenous mobility assumption, we also report the contribution of the job-match effects in

[14] For second-generation immigrants, the pay changes among movers from the top to the bottom are 1.036 log points and the drop for the movers of the opposite direction is 1.008 log points. The changes of the corresponding groups in first-generation immigrants are 0.938 and 1.006 log points (EUR 2.55 and EUR 2.73, respectively). Among refugees, the wages increase from 2.060 to 2.864 log points (EUR 7.85 to EUR 17.53) for upward movers and decrease from 2.756 to 2.132 log points (EUR 15.74 to 8.43) for the downward movers.

[15] Table C.2 in the Appendix presents the movers' wages around the year of transition for each of the groups.

Figure 3.3: Event analysis for exogenous mobility



Note: The figure shows the mean wages of those who changed jobs and held the preceding jobs for two or more years and the new job for two or more years. Each job is classified into quartiles based on the mean log wage of co-workers of both groups in the last year of the old jobs and in the first year of the new job (for the destination firm) by nativity groups.

the next section, and confirm they are not dominant. Both the event analysis and the comparison test allow us to confirm that our sample satisfies the exogenous mobility assumption of the AKM model.

3.4.2 AKM model

Table 3.3 presents the results from Equation (3.2) and shows significant disparities between refugees and natives. The pay gap between natives and refugees is of a larger magnitude than the gap between natives and the two other groups of immigrants. Refugees earn only approximately half of the average hourly wage of native workers. More specifically, on average, native workers earn EUR 17.85 ($e^{2.882}$) per hour whilst the average hourly wage for refugees is EUR 10.40 ($e^{2.342}$). Such a large gap is not observed in the case of the other types of immigrants[16], which confirms that refugees face larger pay disparities as suggested by Brell et al. (2020). Wage variance among refugees is also smaller than among other groups. The standard deviation of log wages in the refugee group is 0.331 compared with a standard deviation between 0.439 and 0.515 for the other groups. This suggests that all refugees earn low hourly wages with small variations and very few refugees achieve high wages.

As mentioned above, we also provide evidence of the exogenous mobility among firms from this estimation. While the RMSE and adjusted R^2 of the models are shown in *Panel A* of Table 3.3, *Panel B* presents the corresponding statistics of the job-match models, which include dummies of person-firm combinations as well as the covariants in the base model. The comparison allows us to check whether the idiosyncratic pay premiums of job-match effects are significant.[17] The job-match component has a standard deviation of 0.093 for natives and 0.078 for refugees, which accounts for less than 5.6% of the overall wage variance for both

[16] Other immigrants earn about 70% of natives' wages in our sample

[17] Large job-match effects may violate the exogenous mobility assumption.

groups. Such a small job-match effect may alleviate concerns about the potential role of job-match effects on mobility patterns (Dostie et al., 2021).

Panel C presents the contribution of each component to the wage variance. It indicates that person effects and covariates are dominant determinants of the wage variance and explain approximately 70% of wage differences within each group. This result is broadly in line with the literature's findings. For instance, Card et al. (2013) and Dostie et al. (2021) report shares of between 63 and 67%. Moreover, the results show that firm effects play a greater role in determining earnings fluctuations among refugees (26.0%) than among natives and other types of immigrants (approximately 10%), and that the covariance between person and firm effects for refugees is strongly negative (-20.7%). The negative covariance indicates that highly paid individuals (i.e. high-ability) are not employed at firms with high pay premiums. In other words, firms are less likely to pay refugees in accordance with their skills thereby inducing a mismatch between abilities and wages, leading to higher firm effects and a negative covariance of person and firm effects. This suggests that high-ability refugees face limited access to higher-paying job opportunities and are under-paid for their level of skills.

Due to data availability, the main dataset has no information on educational attainment. However, since after entering the labour market, educational attainments are relatively constant, educational differences between individuals are likely to be broadly captured by person effects. Indeed, Dostie et al. (2021) compare estimation results with and without education

Table 3.3: Summary of estimated AKM models

	Natives (1)	2nd-gen (2)	1st-gen (3)	Refugee (4)
Mean of log wages	2.882	2.729	2.650	2.342
Standard Deviation of log wages	0.515	0.541	0.439	0.331
Number of person-year obs.	54,555,580	6,710,024	2,281,716	183,628
Panel A: Summary of parameter estimates				
Number of person effects	7,407,587	1,003,001	576,689	48,279
Number of firm effects	513,499	185,902	59,757	16,470
Standard Deviation of person effect	0.392	0.354	0.323	0.279
Standard Deviation of firm effect	0.141	0.159	0.152	0.169
Standard Deviation of $X\beta$	0.249	0.275	0.144	0.389
Standard Deviation of residuals	0.136	0.158	0.139	0.120
Correlation of person and firm effects	0.237	0.143	0.185	-0.241
Correlation of person effect and $X\beta$	-0.135	0.020	0.016	-0.732
Correlation of firm effect and $X\beta$	0.161	0.255	0.142	0.145
Root MSE of model	0.147	0.174	0.163	0.149
Adjusted R^2 of model	0.918	0.897	0.862	0.797
Correlation native/immig firm effect		0.794	0.537	0.411
Panel B: Comparison with job-match effect model				
Number of job-match effects	10,116,653	1,421,298	615,567	49,741
Root MSE of model	0.114	0.132	0.145	0.127
Adjusted R^2 of model	0.947	0.936	0.894	0.846
Standard Deviation of job-match effect	0.093	0.113	0.075	0.078
Panel C: Share of variance of log-wages due to:				
Person effects + $X\beta$	0.713	0.699	0.655	0.641
Firm effects	0.075	0.086	0.119	0.260
Covariance of person and firm effects	0.099	0.055	0.094	-0.207
Covariance of firm and $X\beta$	0.043	0.076	0.032	0.174
Residuals	0.070	0.085	0.099	0.132

Note: The model is estimated separately for natives (Column (1)), second-generation immigrants (Column (2)), first-generation immigrants (Column (3)) and refugees (Column (4)) within the connected set of each group. The job-match effect model includes job-match effects (i.e., dummies for each person-firm combination) as well as other covariates in the basic model. The standard deviation of job-match effect is estimated as the square root of the difference in mean squared errors between the AKM model and job-match effect model.

information, and find similar implications. We also estimate the AKM model with education information, which reduces the sample size and excludes observations from the year 2022. The results presented in Table C.4 are consistent with our main specification, particularly in terms of the contribution of firm effects. For instance, firm effects explain 7.5% of native wage variances based on the main sample without education variables. The equivalent estimate is 7.2% in the sample with education variables. We are therefore confident that the omission of education variables is not unduly influencing our estimates.

3.4.3 Correlations between person and firm effects

As reported in *Panel A* of Table 3.3, the correlations between the estimated person effects and the firm effects (ρ) are positive for all groups (i.e. Natives: 0.237, 2nd-gen: 0.143, 1st-gen: 0.185) except for refugee workers (i.e., -0.241). While several studies, including the original study of Abowd et al. (1999) and the recent study by Aslund et al. (2021), report negative covariance between person and firm effects, Andrews et al. (2008); Card et al. (2018); Dostie et al. (2021) point out that these correlations need to be interpreted with care as they might be biased due to sampling errors. These could arise if networks are weakly connected, i.e. if only a single worker or a small number of workers connect a particular firm to the broader set of firms. In our sample, these sampling errors may be particularly significant in the sample of refugees, the group with the least number of observations in our data and within each firm. The firms in our sample employ, on average, 2.49 refugees compared with 13.14 natives, 8.85 first-generations and 4.65 second-generations.

Gerard et al. (2021) and Dostie et al. (2021) propose to measure assortative matching, which refers to the tendency of the person and firm effects to be matched systematically, to capture the correlations in a different way. For an assortative matching specification, we regress the estimated person effect on the firm effect from the AKM model as below:

$$\widehat{\alpha}_{gi} = \lambda_{0g} + \lambda_{1g}\widehat{\psi}_{J(it)} + \xi_{git} \tag{3.7}$$

Where, λ_{1g} yields a metric to test assortative matching, which correspond to the correlation of the worker and firm effects. We estimate Equation 3.7 with OLS and IV because OLS could be downward biased, again, due to sampling errors. We can use the estimated firm effects for the *other* groups as an instrumental variable for non-biased estimates of λ_{1g}.

Table 3.4 reports the estimation results in the dual-connected set. It shows larger assortative matching in magnitudes in the IV estimation than OLS estimations in the AKM estimations for natives and other immigrants, suggesting the presence of a downward bias. Yet, the small difference between OLS and IV among natives indicates the bias for natives is not substantial, while the bias is larger for immigrants, indicated by the larger gaps between OLS and IV. The assortative matching metric also suggests a negative relationship between person and firm effects among refugees, with both OLS and IV estimations. A firm that pays a 10% higher wage premium has native employees whose average wage capacity is 9.73

percentage points higher, and refugee workers with lower wage capacity by 14.9 percentage points (Columns (2) and (8)). Again, these implications highlight a mismatch between skills and firm-specific premiums, whereas other workers with higher capability obtain higher wage premiums.

Table 3.4: Relationship between estimated person effects and firm effects

	Natives		2nd-gen		1st-gen		Refugees	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
λ_1	0.924***	0.973***	0.376***	0.874***	0.444***	1.428***	-0.430***	-0.149***
	(0.000407)	(0.000524)	(0.000885)	(0.00114)	(0.00145)	(0.00259)	(0.00398)	(0.00763)

Note: Table entries are coefficients from the regression of estimated person effects on estimated firm effects for the firm of an individual's main job in the year. For the IV estimation, the instrumental variable for the estimated firm effect for native workers is the estimated firm effect for immigrant workers (2nd/1st-generation immigrants and refugees) while firm effects for second-generation immigrants are instruments for natives. All estimations are based on the dual-connected sets of workers.

3.4.4 Normalisation of firm effects

To normalise the estimated firm effects from the AKM model, we need to identify the kink point (τ from Equation 3.3) after which firm premiums begin to increase in line with value added per worker. Table C.6 presents the metrics of the goodness of fits of the model with different values of τ, and indicates that 3.8 and 3.75 log points yield the best fit for two combinations: natives and second-generation immigrants; and natives and first-generation immigrants, respectively. When comparing natives and refugees, the threshold with the best fit corresponds to 3.85 log points.

3.4.5 Decomposition into sorting and pay-setting effects

Finally, we use the framework of Equation 3.6 to evaluate the contribution of firm-specific pay premiums to the average wage gap between natives and each of the groups of immigrants, and decompose the total effect into two channels: firm sorting and pay-setting. Table 3.5 presents the results for the overall population of natives and immigrants in each of the dual-connected sets. We present the overall wage gap in Column (1), the mean estimated firm effects for natives and immigrants in Columns (2) and (3), respectively, the total contributions of firm effects to the wage gap in Column (4), the sorting effect in Columns (5), and pay-setting effects in Columns (6).

Table 3.5: Contribution of firm-specific pay premiums to the pay gap

	Pay gap	Mean firm premium		Total contributions	Sorting	Pay-setting
		Native	Immig	of firm effect	effect	effect
	(1)	(2)	(3)	(4)	(5)	(6)
Native - 2nd-gen	0.164	0.057	0.030	0.028	0.005	0.023
				(0.17)	(0.03)	(0.14)
Native - 1st-gen	0.153	0.061	0.034	0.027	-0.004	0.030
				(0.17)	(-0.03)	(0.20)
Native - Refugee	0.325	0.043	0.003	0.040	0.046	-0.007
				(0.12)	(0.14)	(-0.02)

Note: Column (1) shows the mean log wage gap between natives and immigrants, in the dual-connected set. Columns (2) and (3) show the mean normalised pay premiums by groups. Column (4) is the difference between columns (2) and (3), referring to the total contribution of firm-specific sorting and pay-setting policies to the native-immigrant gap. Columns (5) and (6) decompose the total in column (4) into firm sorting effects (5) and pay-setting effects (6). Entries in parentheses are the shares of the overall gap explained by the component in the column.

Column (1) shows that the wage gap between natives and refugees is EUR 4.02 on average[18], almost double the wage gap between natives and other immigrant groups (i.e. EUR 2.26 - 2.44). Even after restricting our sample to the dual-connected set of firms with financial data, which limits the number of native workers in relatively lower-wage firms, we continue to observe a notably larger native-refugee gap than the other combinations, suggesting refugees substantially lag behind other workers. Column (4) shows the total contributions of the firm effects to the entire pay gap, suggesting the firm component is between 0.027 and 0.040 log points across the dual-connected sets. In terms of share to the total pay gaps, the net firm effects account for 12 to 17% across types of immigrants, which is consistent with specifications by Dostie et al. (2021) (i.e. 15 to 22%).

Columns (5) and (6) decompose the total firm effects into sorting and pay-setting effects and show that most of the firm effects can be attributed to pay-setting for gaps between natives and second- or first-generation immigrants (i.e. 14% and 20%, respectively).[19] It indicates that immigrants are paid less than natives in a given firm within the dual-connected firms, and such gaps within firms widen the pay gaps. Notably, sorting effects for first-generation immigrants are negative. This may be explained by the fact that some types of work visas, including those for highly-skilled migrants, set minimum annual earnings as a requirement, which may make immigrants more concentrated within high-paying firms. In contrast, we observe dominant sorting and essentially zero pay-setting effects for the native-refugee gap.

[18] EUR 14.49 / 2.673 log points (native) - EUR 10.46 / 2.348 log points

[19] Whereas Card et al. (2016); Dostie et al. (2021) report larger sorting effects than pay-setting effect, Li et al. (2023) observe larger pay-setting effects for younger workers. Within age groups younger than 30, pay-setting effects explain 19% of gender pay gaps in Canada, albeit sorting accounts for approximately 10%.

In other words, once refugees get a job, they are paid equally to native co-workers and face no discrimination within firms, whilst the other immigrants experience different wages from natives to some extent. Yet, a high share of firm sorting implies that refugees are unable to access firms with higher pay premiums. The almost zero impact of pay-setting on the native-refugee pay gap may be because refugees are employed at a limited number of firms, offering lower and less heterogenous wages for workers in any group.

Comparing the pay gap in the dual-connected sets with those in the connected sets, we observe considerably smaller gaps in the dual-connected sets. For example, the gap between natives and first-generation immigrants in the connected sets is EUR 3.70 [20], yet, that in the dual-connected set shrinks to EUR 2.26.[21] Likewise, the native-refugee gap in the connected set is EUR 7.44[22], but, declines to EUR 4.02 in the dual-connected set. The differences in the gaps between the connected set and the dual-connected set are also attributed to sorting because focusing on the dual-connected sets drops the firms that employ only either natives or immigrants. Indeed, summary statistics also indicate that the average wages of natives in the connected sets are approximately EUR 20, whilst the corresponding wages in the dual-connected with first-generation immigrants or refugees decrease to EUR 16-18 (Tables 3.1 and C.1). This also reinforces the point that first-generation immigrants and refugees have no access to higher-paying job opportunities.

[20] i.e., EUR 17.85 / 2.882 log points (native) - EUR 14.15 / 2.650 log points (first-generation immigrants) from Table 3.1

[21] EUR 15.98 / 2.771 log points - EUR 13.72 / 2.619 log points

[22] In the connected set, EUR 17.85 / 2.882 log points - EUR 10.40 / 2.342 log points

3.4.6 Robustness check - normalising with sales per worker

To normalise firm premiums, we use value added per worker as a measurement of firm productivity. Dostie et al. (2021) acknowledge that normalising firm effects could affect the values of the pay-setting effects. Following Card et al. (2016), we also use sales per worker as an alternative measure of firm productivity.[23]

For normalisation, the thresholds for low- and high-productivity are 5.4 log points for dual-connected sets between natives and second- and first-generation immigrants, and 5.5 for the sets between natives and refugees (Table C.9). Table C.10 presents the decomposition results based on normalised firm premiums with sales per worker. Our main findings are confirmed by this robustness test.

3.5 Conclusion and discussions

Our study investigates whether and to what extent firm sorting and pay setting effects influence the native-refugee wage gap, using Dutch employer-employee data from 2014 to 2022. We apply the AKM model to quantify the role of firm effects in explaining wage

[23] While Card et al. (2016) exclude particularly low productivity industries, hotels and restaurants, we similarly exclude the hospitality industry, which contains a high proportion of lower productivity firms (i.e. below the threshold) as shown in Table C.8

disparities *within* groups, and a Blinder-Oxaca style decomposition model to specify the contribution of firm-sorting and pay-setting effects to wage gaps *between* groups.

From the AKM estimation, we find that refugees substantially lag behind not only natives but other types of immigrants, with the wage variance among refugees also being smaller than for the other groups. Such a small wage variance among refugees implies that refugees' wage distribution does not contain a long upper tail, in other words, no refugees receive exceptionally high wages, unlike the other groups. We also reveal a negative relationship between person and firm effects for refugee workers, suggesting firms are less likely to pay refugees in accordance with their skills thereby inducing a mismatch between ability and wages, leading to higher firm effects and a negative covariance of person and firm effects. In other words, high-ability refugees face limited access to higher-paying job opportunities, which may result in significant underpay as a whole.

Furthermore, we find that firm-specific pay premiums, the net contribution of sorting and pay-setting effects, form a non-negligible component of wage gaps *within* and *between* groups. However individual (un)observed characteristics, captured by person effects and time-varying covariants, are the main driver of wage disparities, which is consistent with the previous literature. Sorting affects account for most of the contribution of firm-specific pay premiums to the entire wage gaps while pay-setting effects make a minimal contribution to the native-refugee gap, suggesting that refugees are almost paid equally to native co-workers once they get a job. Therefore, our estimation suggests that the large pay gap between natives and refugees are attributed to individual characteristics, and firm sorting effects, rather than pay setting *within* firms. The adverse impact of firm sorting may result from refugees having limited access to higher-paying firms, leading to a mismatch between refugees' abilities and their compensation. To facilitate the economic assimilation of refugees, it is crucial for higher-paying firms to provide accessible job opportunities based on a fair assessment of refugees' skills, thus enabling them to progress through the job ladder towards firms with greater pay premiums (Arellano-Bover and San, 2020).

In order to compare wages between natives and immigrants, it is necessary to restrict our sample to the dual-connected sets for the purpose of decomposition identification. However, this restriction excludes native-dominant firms, which are expected to offer higher wages, thus overlooking the total pay disparities that exist in these firms. Card et al. (2016), whose methodology this study is based on, examines gender pay gaps and the number of male-dominant firms is likely to be limited. As a result, for Card et al. (2016), focusing on the dual-connected set of both genders does not result in a significant loss of sample or notable changes in the characteristics of the decomposition sample. However, we find that this is not the case in our analysis of the native-immigrant wage gap due to the limited number of immigrants across all categories. Consequently, it is crucial for future studies to consider this limitation when conducting decomposition estimations.

Conclusion

This doctoral thesis extensively examines the determinants and obstacles that influence the legal and economic integration of refugees who arrived in the Netherlands in 2014 or later, utilising administrative data. The research findings shed light on several crucial aspects of refugee integration from which meaningful policy implications are derived.

Chapter 1 focuses on legal integration and identifies the level of urbanisation and co-national density in the assignment location as important determinants. Refugees residing in densely populated areas demonstrate a higher likelihood of successfully passing the civic integration exam and eventually acquiring citizenship. However, the marginal effect of urbanisation on integration declines as urbanisation increases and becomes negative at high levels of urban density. It is further shown that co-national density is an additional mechanism that affects integration, indicating that the presence of co-nationals increases social contacts, eases adaptation to the host country and enhances the transmission and acquisition of knowledge. The results indicate that individual factors also reduce the likelihood of refugees achieving legal integration. These factors include suffering from chronic physical or mental health conditions, as well as incurring lengthy waiting times for asylum decisions. The positive impacts of being in employment and being educated within the Netherlands are also identified. These results provide a number of insights for policy. First, if it is difficult to avoid locating refugees in lower-density areas, additional support may be required to ensure that refugees are not missing out on opportunities for knowledge diffusion, networking, and employment. Second, providing refugees with employment and educational opportunities will have a clear benefit to their prospects for legal integration, gaining human capital and working experience. Third, increasing awareness of and access to appropriate health services for refugees and additional resourcing of the asylum process are likely to yield wider and longer-term societal benefits beyond the utility of the individual refugees.

Chapter 2 examines the impact of the introduction of the Orientation on the Dutch Labour Market (ONA) component of the Dutch civic integration exam on the economic performance of refugees in the Netherlands. The chapter's main findings indicate that while the ONA increased the probability of employment for refugees, it did not lead to an improvement in their job quality. The divergence in employment probability between those who took the ONA and those who did not grew throughout the analysis period, reaching 6.4-11.0 percentage points after eight years. But no evidence is found to suggest that the ONA influenced refugees' working hours or hourly wages. It is also found that the benefits of the ONA were predominantly observed among male refugees and those from countries other than Syria. Based on these findings, there are three key conclusions for policymakers. Firstly, civic integration requirements do contribute to economic integration among refugees. Secondly, a dedicated requirement focusing on the demonstration of the active use of host-country-specific knowledge, as required by the ONA, can enhance refugees' ability to participate in labour markets. Thirdly, however, while such a training programme may improve employment prospects, it does not seem to address the issue of low job quality for refugees. The findings of this chapter suggest that other factors, such as skills mismatch or labour market discrimination, may continue to limit the types of jobs accessible to refugees. In light of these findings, policymakers should consider addressing these specific challenges to create more inclusive labour markets while exploring strategies to bridge the gap between employment rates and job quality for all refugees.

Chapter 3, investigates whether and to what extent firm sorting influences the native-refugee wage gap. It does so by applying the AKM model to quantify the role of firm effects in explaining wage disparities *within* groups, and a Blinder-Oxaca style decomposition model to specify the contribution of firm-sorting and pay-setting effects to wage gaps *between* groups. It is found that firm-specific pay premiums, the net contribution of sorting and pay-setting effects, form a non-negligible component of wage gaps *within* and *between* groups. Nevertheless, individual (un)observed characteristics, captured by person effects and time-varying covariants, are the main driver of wage disparities. Sorting effects account for most of the contribution of firm-specific pay premiums to the overall wage gap, while pay-setting effects make a minimal contribution to the native-refugee gap, suggesting that refugees are almost paid equally to native co-workers once they do get a job. Therefore, these results suggest that the large pay gap between natives and refugees is attributed to individual characteristics and firm sorting effects rather than pay setting *within* firms. The adverse impact of firm sorting may result from refugees having limited access to higher-paying firms, leading to a mismatch between their abilities and compensation. To facilitate refugee economic assimilation, it is crucial for higher-paying firms to provide accessible job opportunities based on a fair assessment of refugees' skills, thus enabling them to progress through the job ladder towards firms with greater pay premiums.

For future research, it is essential to explore country comparisons across Europe since refugee integration is a common and urgent issue for European countries. Even though we confirm the Netherlands is representative of the whole of Europe in terms of the number of asylum applications per capita, it obviously remains distinct in terms of asylum/integration programmes and restrictions, local labour market conditions, co-national stock, and perhaps, the public attitude toward refugees across Europe. However, this attempt brings a critical challenge. The main restriction has been the absence of rich and relevant data, which allows us to observe individual residential status, labour market outcomes, and asylum-related administration information. Potentially, the European Labour Force Survey would be an option for international comparisons, yet, it has limited information, in particular, on asylum administration and in a repeated cross-sectional structure. The improvement of the micro-level data resources would address these research challenges.

Bibliography

Abowd, J. M., Creecy, R. H., and Kramarz, F. (2002). Computing person and firm effects using linked longitudinal employer-employee data. *Longitudinal Employer-Household Dynamics Technical Papers 2002-06*.

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, 67(2):251–333.

Adamson, G. (2020). Migrants and crime in Sweden in the twenty-first century. *Society*, 57(1):9–21.

Ahrens, A., Beerli, A., Hangartner, D., Kurer, S., and Siegenthaler, M. (2023). The labor market effects of restricting refugees' employment opportunities.

Aksoy, C. G. and Poutvaara, P. (2021). Refugees' and irregular migrants' self-selection into Europe. *Journal of Development Economics*, 152:102681.

Albert, C., Glitz, A., and Llull, J. (2021). Labor market competition and the assimilation of immigrants. *CEPR Discussion Paper*, DP16432.

Albrecht, C., Pérez, M. H., and Stitteneder, T. (2021). The integration challenges of female refugees and migrants: Where do we stand? In *CESifo Forum*, volume 22, pages 39–46. München: ifo Institut-Leibniz-Institut für Wirtschaftsforschung an der Universität München.

Algemene Rekenkamer (2017). *Inburgering: Eerste resultaten van de Wet inburgering 2013* [Integration: First results of the 2013 integration act]. Available at: `https://www.rekenkamer.nl/publicaties/rapporten/2017/01/24/inburgering` (Accessed on 22 September 2021).

Andrews, M. J., Gill, L., Schank, T., and Upward, R. (2008). High wage workers and low wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(3):673–697.

Arellano-Bover, J. and San, S. (2020). The role of firms in the assimilation of immigrants. *Available at SSRN 3594778*.

Arendt, J. N. (2022). Labor market effects of a work-first policy for refugees. *Journal of Population Economics*, 35(1):169–196.

Arendt, J. N., Dustmann, C., and Ku, H. (2022). Refugee migration and the labour market: lessons from 40 years of post-arrival policies in Denmark. *Oxford Review of Economic Policy*, 38(3):531–556.

Aslund, O., Bratu, C., Lombardi, S., and Thoresson, A. (2021). Firm productivity and immigrant-native earnings disparity. *IZA Discussion Paper*, 14960.

Aslund, O., Engdahl, M., Rosenqvist, O., et al. (2022). Limbo or leverage? Asylum waiting and refugee integration. *IZA Discussion Paper*, 15063.

Auer, D. (2018). Language roulette–the effect of random placement on refugees' labour market integration. *Journal of Ethnic and Migration Studies*, 44(3):341–362.

Bahar, D. and Rapoport, H. (2018). Migration, knowledge diffusion and the comparative advantage of nations. *The Economic Journal*, 128(612):F273–F305.

Battisti, M., Giesing, Y., and Laurentsyeva, N. (2019). Can job search assistance improve the labour market integration of refugees? Evidence from a field experiment. *Labour Economics*, 61:101745.

Battu, H., Seaman, P., and Zenou, Y. (2011). Job contact networks and the ethnic minorities. *Labour Economics*, 18(1):48–56.

Baum, C. F., Lööf, H., Stephan, A., and Zimmermann, K. F. (2020). Occupational sorting and wage gaps of refugees. *Life Course Centre Working Paper*, (2020-14).

Beaman, L. A. (2012). Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the US. *The Review of Economic Studies*, 79(1):128–161.

Bertoli, S., Brücker, H., and Fernández-Huertas Moraga, J. (2020). Do processing times affect the distribution of asylum seekers across Europe? *IZA Discussion Papers*, 13018.

Bevelander, P. (2020). Integrating refugees into labor markets: Economic integration of refugees into their host country is important and benefits both parties. *IZA World of Labor*.

Bevelander, P. and Veenman, J. (2006). Naturalisation and socioeconomic integration: The case of the Netherlands. *IZA Discussion Papers*, 2153.

Blackwell, M., Iacus, S., King, G., and Porro, G. (2009). CEM: Coarsened exact matching in Stata. *The Stata Journal*, 9(4):524–546.

Boeri, T., De Philippis, M., Patacchini, E., and Pellizzari, M. (2015). Immigration, housing discrimination and employment. *The Economic Journal*, 125(586):F82–F114.

Borjas, G. J. (2015). The slowdown in the economic assimilation of immigrants: Aging and cohort effects revisited again. *Journal of Human Capital*, 9(4):483–517.

Borjas, George, J. (2020). *Education*. McGraw-Hill Education.

Braun, S. T. and Dwenger, N. (2020). Settlement location shapes the integration of forced migrants: Evidence from post-war Germany. *Explorations in Economic History*, 77:101330.

Brell, C., Dustmann, C., and Preston, I. (2020). The labor market integration of refugee migrants in high-income countries. *Journal of Economic Perspectives*, 34(1):94–121.

Broom NT2 (n.d.). What is orientation on the Dutch labour market (ONA)? Available at: https://www.nt2.nl/en/dossier/orientatie_op_de_nederlandse_arbeidsmarkt/wat_is_orientatie_op_de_nederlandse_arbeidsmarkt (Accessed on 21 September 2021).

Brücker, H., Glitz, A., Lerche, A., and Romiti, A. (2021). Occupational recognition and immigrant labor market outcomes. *Journal of Labor Economics*, 39(2):497–525.

Caliendo, M. and Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1):31–72.

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2019). Regression discontinuity designs using covariates. *Review of Economics and Statistics*, 101(3):442–451.

Card, D. (2009). Immigration and inequality. *American Economic Review*, 99(2):1–21.

Card, D., Cardoso, A. R., Heining, J., and Kline, P. (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics*, 36(S1):S13–S70.

Card, D., Cardoso, A. R., and Kline, P. (2016). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly journal of economics*, 131(2):633–686.

Card, D., Heining, J., and Kline, P. (2013). Workplace heterogeneity and the rise of West German wage inequality. *The Quarterly journal of economics*, 128(3):967–1015.

Cardoso, A. R., Guimaraes, P., and Portugal, P. (2016). What drives the gender wage gap? a look at the role of firm and job-title heterogeneity. *Oxford Economic Papers*, 68(2):506–524.

CBS (2019). *Asiel en integratie: Cohortonderzoek asielzoekers en statushouders 2019* [Asylum and integration: Cohort survey asylum seekers and status holders 2019]. Available at: https://longreads.cbs.nl/asielenintegratie-2019/ (Accessed on 6 August 2020).

Chung, M. C., AlQarni, N., AlMazrouei, M., Al Muhairi, S., Shakra, M., Mitchell, B., Al Mazrouei, S., and Al Hashimi, S. (2018). The impact of trauma exposure characteristics on post-traumatic stress disorder and psychiatric co-morbidity among Syrian refugees. *Psychiatry Research*, 259:310–315.

COA (n.d.). Reception centre and amenities. Available at: https://www.mycoa.nl/en/infosheetprint/1553 (Accessed on 31 May 2021).

Crawley, H. and Hagen-Zanker, J. (2019). Deciding where to go: Policies, people and perceptions shaping destination preferences. *International Migration*, 57(1):20–35.

Dagnelie, O., Mayda, A. M., and Maystadt, J. F. (2019). The labor market integration of refugees in the United States: Do entrepreneurs in the network help? *European Economic*

Review, 111:257–272.

Dahlberg, M., Egebark, J., Özcan, G., and Vikman, U. (2022). Labor market integration of refugees: RCT evidence from an early intervention program in Sweden. *arXiv preprint arXiv:2203.00487*.

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

De La Roca, J. and Puga, D. (2017). Learning by working in big cities. *The Review of Economic Studies*, 84(1):106–142.

De Vroome, T., Coenders, M., Van Tubergen, F., and Verkuyten, M. (2011). Economic participation and national self-identification of refugees in the Netherlands. *International migration review*, 45(3):615–638.

De Vroome, T. and Van Tubergen, F. (2010). The employment experience of refugees in the Netherlands. *International Migration Review*, 44(2):376–403.

Department of Dutch, UCL (n.d.). Dutch government and the immigrants. Available at: `https://www.ucl.ac.uk/dutchstudies/an/SP_LINKS_UCL_POPUP/SPs_english/multicultureel_gev_ENG/pages/autochtonen.html` (Accessed on 29 November 2021).

Dostie, B., Li, J., Card, D., and Parent, D. (2021). Employer policies and the immigrant-native earnings gap. *Journal of Econometrics*, 233(2):544–567.

DUO (2022). *Voortgangscijfers Inburgering Totaal Nederland* [Progress figures integration: Total Netherlands]. Available at: `https://duo.nl/zakelijk/inburgering-ketenpartners/cijfers-over-inburgeren/cijfers-inburgeren.jsp` (Accessed on 30 June 2022).

Duranton, G. and Puga, D. (2004). Micro-foundations of Urban Agglomeration Economies. In *Handbook of Regional and Urban Economics*, volume 4 of *Handbook of Regional and Urban Economics*, pages 2063–2117. Elsevier.

Dutch Central Government (n.d.). Civic integration act 2021. Available at: `https://www.rijksoverheid.nl/onderwerpen/inburgeren-in-nederland/nieuwe-wet-inburgering` (Accessed on 22 October 2021).

Eckert, F., Hejlesen, M., and Walsh, C. (2022). The return to big-city experience: Evidence from refugees in Denmark: The return to big-city experience. *Journal of Urban Economics*, 130:103454.

Edin, P.-A., Fredriksson, P., and Åslund, O. (2004). Settlement policies and the economic success of immigrants. *Journal of Population Economics*, 17(1):133–155.

Eeckhout, J. and Kircher, P. (2011). Identifying sorting—in theory. *The Review of Economic Studies*, 78(3):872–906.

Enchautegui, M. E. and Giannarelli, L. (2015). The economic impact of naturalization on immigrants and cities. *Urban Inst*, pages 18–24.

Engdahl, M., Lindgren, K.-O., and Rosenqvist, O. (2020). The role of local voting rights for non-naturalized immigrants: a catalyst for integration? *International Migration Review*, 54(4):1134–1157.

Ersanilli, E. (2014). focus MIGRATION: Country profile Netherlands. *Institute for Migration Research and Intercultural Studies (IMIS)*.

Eurostat (2021). Data browser: Asylum and first time asylum applicants. `https://ec.europa.eu/eurostat/databrowser/view/migr_asyappctza/default/table?lang=en` (Accessed on 4 February 2021).

Fadhlia, T. N., Sauter, D. A., and Doosje, B. (2022). Adversity, emotion, and resilience among Syrian refugees in the Netherlands. *BMC psychology*, 10(1):1–13.

Fasani, F., Frattini, T., and Minale, L. (2021). Lift the ban? initial employment restrictions and refugee labour market outcomes. *Journal of the European Economic Association*, 19(5):2803–2854.

Fasani, F., Frattini, T., and Minale, L. (2022). (The struggle for) refugee integration into the labour market: Evidence from Europe. *Journal of Economic Geography*, 22(2):351–393.

Fiolet Taaltrainingen (n.d.). What are the costs of the civic integration course? Available at: `https://fiolet.nl/civic-integration-course/costs-civic-integration-course/` (Accessed on 28 February 2022).

Foged, M., Hasager, L., Peri, G., Arendt, J. N., and Bolvig, I. (2022). Language training and refugees' integration. *The Review of Economics and Statistics*, pages 1–41.

Foged, M. and Werf, C. V. D. (2022). Access to Language Training and Local Integration of Refugees. *CReAM Discussion Paper Series*, (April).

Garrido, M. M., Kelley, A. S., Paris, J., Roza, K., Meier, D. E., Morrison, R. S., and Aldridge, M. D. (2014). Methods for constructing and assessing propensity scores. *Health Services Research*, 49(5):1701–1720.

Gathmann, C. and Keller, N. (2018). Access to citizenship and the economic assimilation of immigrants. *The Economic Journal*, 128(616):3141–3181.

Gerard, F., Lagos, L., Severnini, E., and Card, D. (2021). Assortative matching or exclusionary hiring? the impact of employment and pay policies on racial wage differences in Brazil. *American Economic Review*, 111(10):3418–57.

Godøy, A. (2017). Local labor markets and earnings of refugee immigrants. *Empirical Economics*, 52(1):31–58.

Goodman, S. W. and Wright, M. (2015). Does mandatory integration matter? Effects of civic requirements on immigrant socio-economic and political outcomes. *Journal of Ethnic and Migration Studies*, 41(12):1885–1908.

Govind, Y. (2021). Is naturalization a passport for better labor market integration? Evidence from a quasi-experimental setting.

Hainmueller, J., Hangartner, D., and Lawrence, D. (2016). When lives are put on hold: Lengthy asylum processes decrease employment among refugees. *Science Advances*, 2(8):e1600432.

Hainmueller, J., Hangartner, D., and Pietrantuono, G. (2017). Catalyst or crown: Does naturalization promote the long-term social integration of immigrants? *American Political Science Review*, 111(2):256–276.

Hainmueller, J., Hangartner, D., and Ward, D. (2019). The effect of citizenship on the long-term earnings of marginalized immigrants: Quasi-experimental evidence from Switzerland. *Science Advances*, 5(12):eaay1610.

Helgesson, M., Wang, M., Niederkrotenthaler, T., Saboonchi, F., and Mittendorfer-Rutz, E. (2019). Labour market marginalisation among refugees from different countries of birth: a prospective cohort study on refugees to Sweden. *J Epidemiol Community Health*, 73(5):407–415.

Henskens, B. and Dijsselbloem, J. (2020). Revisiting the 2013 Dutch integration law. Available at: `https://www.diggitmagazine.com/articles/dutch-integration-Netherlands` (Accessed on 17 September 2021).

Hoon, M. d., Vink, M., and Schmeets, H. (2019). A ticket to mobility? naturalisation and subsequent migration of refugees after obtaining asylum in the Netherlands. *Journal of Ethnic and Migration Studies*, pages 1–20.

Huddleston, T. and Vink, M. P. (2015). Full membership or equal rights?: The link between naturalisation and integration policies for immigrants in 29 European states. *Comparative Migration Studies*, 3(1):1–19.

Hunt, V., Layton, D., and Prince, S. (2015). Diversity matters. *McKinsey & Company*, 1(1):15–29.

IND (2019). *Hoe kunt u Nederlander worden?* [How can you become a Dutch citizen?]. Available at: `https://ind.nl/en/dutch-citizenship/Pages/Naturalisation.aspx` (Accessed on 5 May 2020).

IND (n.d.). Civic integration for more secure residence permit and naturalisation. Available at: `https://ind.nl/en/Pages/Integration-in-the-netherlands.aspx` (Accessed on 22 October 2021).

IOM (2019). World migration report 2020.

Klaesson, J., Öner, Ö., and Pennerstorfer, D. (2021). Getting the first job: Size and quality of ethnic enclaves and refugee labor market entry. *Journal of Regional Science*, 61(1):112–139.

Kone, Z., Ruiz, I., and Vargas-Silva, C. (2019). Refugees and the UK labour market. *Compas*

Report.

Koster, H. R. A. and Ozgen, C. (2021). Cities and Tasks. *Journal of Urban Economics*, 126(103386).

Language Centre, University of Groningen (n.d.). Orientation to the Dutch labour market (ONA). Available at: https://www.rug.nl/language-centre/language-courses/dutch/civic-integration/orientatie-nederlandse-arbeidsmarkt?lang=en (Accessed on 10 January 2022).

Lazear, E. P. and Gibbs, M. (2014). *Personnel economics in practice - Third edition*. John Wiley & Sons.

Li, J., Dostie, B., and Simard-Duplain, G. (2023). Firm pay policies and the gender earnings gap: The mediating role of marital and family status. *ILR Review*, 76(1):160–188.

Li, M. (2013). Using the propensity score method to estimate causal effects: A review and practical guide. *Organizational Research Methods*, 16(2):188–226.

Lochmann, A., Rapoport, H., and Speciale, B. (2019). The effect of language training on immigrants' economic integration: Empirical evidence from France. *European Economic Review*, 113:265–296.

Lochner, B. and Schulz, B. (2021). Firm productivity, wages, and sorting. *Journal of Labour Economics*.

Manning, A. and Swaffield, J. (2008). The gender gap in early-career wage growth. *The Economic Journal*, 118(530):983–1024.

Martén, L., Hainmueller, J., and Hangartner, D. (2019). Ethnic networks can foster the economic integration of refugees. *Proceedings of the National Academy of Sciences*, 116(33):16280–16285.

McKenzie, D. and Rapoport, H. (2010). Self-selection patterns in Mexico-US migration. *the Review of Economics and Statistics*, 92(4):811–821.

Mossaad, N., Ferwerda, J., Lawrence, D., Weinstein, J. M., and Hainmueller, J. (2018). Determinants of refugee naturalization in the United States. *Proceedings of the National Academy of Sciences*, 115(37):9175–9180.

Müller, T., Pannatier, P., and Viarengo, M. (2023). Labor market integration, local conditions and inequalities: Evidence from refugees in Switzerland. *World Development*, 170:106288.

NL TIMES (2016). Netherlands counts 64,000 registered Syrian residents, 28 september, 2016. Available at: https://nltimes.nl/2016/09/28/netherlands-counts-64000-registered-syrian-residents (Accessed on 30 May 2021).

Norton, E. C. (2022). The inverse hyperbolic sine transformation and retransformed marginal effects. *The Stata Journal*, 22(3):702–712.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, pages 693–709.

OECD (2019a). Education at a glance: OECD indicators. Available at: https://www.oecd.org/education/education-at-a-glance/EAG2019_CN_NLD.pdf (Accessed on 26 July 2023).

OECD (2019b). International migration statistics: International migration database. Available at: https://data.oecd.org/migration/foreign-born-population.htm (Accessed on 30 January 2021).

OECD (2019c). Ready to help?: Improving resilience of integration system for refugee and other vulnerable migrants.

Ozgen, C. (2021). The economics of diversity: Innovation, productivity and the labour market. *Journal of Economic Surveys*, 35(4):1168–1216.

Ozgen, C., Nijkamp, P., and Poot, J. (2013). The impact of cultural diversity on firm innovation: evidence from Dutch micro-data. *IZA Journal of Migration*, 2(1):18.

Patacchini, E. and Zenou, Y. (2012). Ethnic networks and employment outcomes. *Regional Science and Urban Economics*, 42(6):938–949.

Portes, A. and Curtis, J. W. (1987). Changing flags: Naturalization and its determinants among Mexican immigrants. *International Migration Review*, 21(2):352–371.

Putnam, R. D. (2007). E pluribus unum: Diversity and community in the twenty-first century the 2006 Johan Skytte Prize lecture. *Scandinavian political studies*, 30(2):137–174.

Rhoades, S. A. (1993). The herfindahl-hirschman index. *Fed. Res. Bull.*, 79:188.

Robinson, V. and Segrott, J. (2002). *Understanding the decision-making of asylum seekers*, volume 12. Home Office London.

Sarvimäki, M. and Hämäläinen, K. (2016). Integrating immigrants: The impact of restructuring active labor market programs. *Journal of Labor Economics*, 34(2):479–508.

Stichting Blik op Werk (n.d.). Civic integration - find a good school nearby. Available at: https://www.zoekinburgerschool.nl/ (Accessed on 11 July 2022).

Stips, F. and Kis-Katos, K. (2020). The impact of co-national networks on asylum seekers' employment: Quasi-experimental evidence from Germany. *PloS one*, 15(8):e0236996.

Tolsma, J., Lameris, J., and Savelkoul, M. (2021). Exposure to asylum seekers and changing support for the radical right: A natural experiment in the Netherlands. *PLoS One*, 16(2):e0245644.

UNHCR (2021). Global trends in forced displacement 2020. Available at: https://www.unhcr.org/uk/statistics/unhcrstats/60b638e37/global-trends-forced-displacement-2020.html (Accessed on 22

September 2021).

UNHCR (2023). Global trends forced displacement 2022. Available at: https://www.unhcr.org/sites/default/files/2023-06/global-trends-report-2022.pdf (Accessed on 23 July 2023).

DUO (n.d.). *Inburgeren*. Available at: https://www.inburgeren.nl/en/index.jsp (Accessed on 27 September 2021).

Varela, V. V., Suárez, B. F., and DePalma, R. (2020). What is integration? a comparative view from immigrants and municipal policy. *International Migration*.

World Bank (2021). World development indicators - total population. Available at: https://data.worldbank.org/indicator/SP.POP.TOTL?locations=EU (Accessed on 2 December 2021).

World Bank (2022). Labor force participation rate, female - Syrian Arab Republic. Available at: https://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS?locations=SY (Accessed on 1 February 2023).

Yang, P. Q. (1994). Explaining immigrant naturalization. *International migration review*, 28(3):449–477.

Yao, Y. and van Ours, J. C. (2015). Language skills and labor market performance of immigrants in the Netherlands. *Labour Economics*, 34:76–85.

Yumoto, H., Cole, M., Jabbour, L., and Ozgen, C. (2023a). Invisible immigrants: Urbanisation, co-national density and the legal integration of refugees. *Working paper*.

Yumoto, H., Cole, M., Jabbour, L., and Ozgen, C. (2023b). Local labour market training and civic integration exams: Examining the employment outcomes of refugees. *Working paper*.

Appendix A

Appendix for Chapter 1

A.1 Variable descriptions

Table A.1: List of dependent variables

No.	Variable	Description
1	Pass Exam	An indicator taking the value of 1 if the individual passed the civic integration exam by 2020, 0 otherwise.
2	Naturalisation	An indicator that the individual is naturalised, taking a value 1 if the individual is recorded as Dutch in municipal registers by 2020, 0 otherwise.

Table A.2: List of independent variables at the individual level

No.	Variable	Description
1	Years since migration (YSM)	A variable that indicates the number of years that the individual has spent in the Netherlands since migration
2	Arrival year	Dummy variables that indicate when an individual arrived in the Netherlands.
3	Arrival age (Age at arrival)	Age when the individual registered at COA refugee centres in the Netherlands for the first time
4	Female	Equal to 1 if the individual is female, 0 otherwise.
5	Education in the Netherlands	The indicator that the individual has enrolled in educational programmes in the Netherlands: Low-Middle or High
6	Family size	The number of individuals in the asylum application
7	Position in household	An indicator of the individual's position in the household. This includes Single; living with a married or unmarried partner and no children (Spouse/Partner); living with children regardless of their marital status (Parent); and living with their parent (Child).
8	Country of origin	The country of origin of each individual. The 5 largest refugee countries of origin (i.e. Syria; Eritrea; Ethiopia; Iran: and Iraq) are coded as dummies for each and the remainder are assigned to Other.
9	Stateless	A dummy indicating the individual is stateless.
10	Asylum application months	Number of months since the individual arrived in the Netherlands to be granted an asylum residence permit
11	Mental health problems	A dummy variable indicating if the individual undertook a mental health consultation within the health insurance during the period
12	Chronic physical problems	A dummy variable indicating if the individual took physiotherapy or similar within health insurance during the period
13	Duration of social benefits	Share of months that the individual received social assistance benefits, which include the Work and Social Assistance Act (WWB) in the period, out of the period in the Netherlands
14	Currently employed	A dummy variable indicating whether the individual is employed either at the moment of passing the exam (for *pass exam*), or in 2020 (for *naturalisation*)
15	Hours worked in urban	Accumulated number of hours worked in urban areas, in which municipalities categorised into 'Very strong' and 'Strong' urban, until the exam or by 2020
16	Hours worked in rural	Accumulated number of hours worked in rural areas, in which municipalities categorised into 'Moderate', 'Weak' and 'Non-urban', until the exam or by 2020

Table A.3: List of independent variables at the municipality level

No.	Variable	Description
1	Population density	A number of people per square kilometre of land. In this paper, it is shown in 1000 scale (1000 people/km^2).
2	Co-national density	A number of people with the same group of nationalities per square kilometre of land. In this paper, it is shown in 1000 scale (1000 people/km^2). The nationality groups are based on the Standard country or area codes for statistical use (M49) by the United Nations Statistics Division (UNSD). In the case of stateless, it is assumed no co-national population.
3	Employed co-national density	A number of employed people with the same group of nationalities per square kilometre of land. In this paper, it is shown in 1000 scale (1000 people/km^2). The categories and an assumption of no co-nationality populations for the stateless are the same as co-nationality density.
4	Co-national stock	A number of people with the same group of nationalities in municipalities. In this paper, it is shown on 1000 scale (1000 people/municipality). The categories and an assumption of no co-nationality populations for the stateless are the same as co-nationality density.
5	Unemployment rate	The unemployment rate of each municipality out of the total population aged 15 to 75, taking a value between 0 to 1.

A.2 Summary statistics

Table A.4: Summary Statistics (Cohort 2014)

	(1) Full sample		(2) Not integrated		(3) Passed exam		(4) Naturalised	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Passed exam/Exempted	89.28%							
Naturalised	75.01%							
Years since migration	5.96	0.23	5.97	0.25	5.95	0.23	5.95	0.23
Asylum application mon.	4.02	3.60	5.25	6.46	3.88	3.05	3.70	3.03
Age at arrival	30.78	11.08	28.79	18.49	31.02	9.80	31.73	10.52
Female	22.62%		27.15%		22.07%		22.66%	
Education in NL								
No education	82.43%		59.14%		85.22%		84.55%	
Low or middle	15.32%		38.58%		12.53%		12.83%	
High	2.25%		2.28%		2.25%		2.62%	
Country of origin								
Syria	59.23%		42.71%		61.21%		67.68%	
Ethiopia	8.38%		6.94%		8.55%		6.77%	
Eritrea	16.42%		21.18%		15.85%		11.52%	
Iraq	2.45%		6.50%		1.96%		2.15%	
Iran	1.77%		2.46%		1.69%		1.48%	
Stateless	14.11%		13.27%		14.21%		17.64%	
Economic conditions								
Duration of social ben.	45.33%	17.18%	33.96%	33.86%	46.70%	13.25%	44.97%	14.64%
Currently employed	40.57%		33.22%		41.46%		41.31%	
Hours worked in urban	0.88	1.56	0.67	1.32	0.91	1.58	0.89	1.59
Hours worked in rural	0.59	1.39	0.31	0.92	0.62	1.43	0.62	1.43
Municipal characteristics								
Unemployment rate	3.92%	0.86%	4.05%	0.88%	3.90%	0.86%	3.87%	0.85%
Population density	1.89	1.79	2.02	1.82	1.88	1.79	1.83	1.75
Co-national density	0.04	0.08	0.03	0.07	0.04	0.08	0.04	0.08
Employed co-nat. den.	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02
Co-national stock	4.51	11.75	4.26	11.63	4.54	11.77	4.60	11.82
Health conditions								
Mental health problems	6.74%		8.79%		6.50%		5.96%	
Chronic physical prob.	1.32%		4.75%		0.91%		1.22%	
N	10,620		1,138		9,482		7,966	

Note: Column (2) presents who did not pass the civic integration exam nor naturalised. Column (3) presents those who passed the civic integration exam or were exempted. Hours worked in urban/rural, population density, co-national density, employed co-national density, and co-national stock are shown in 1000 scale.

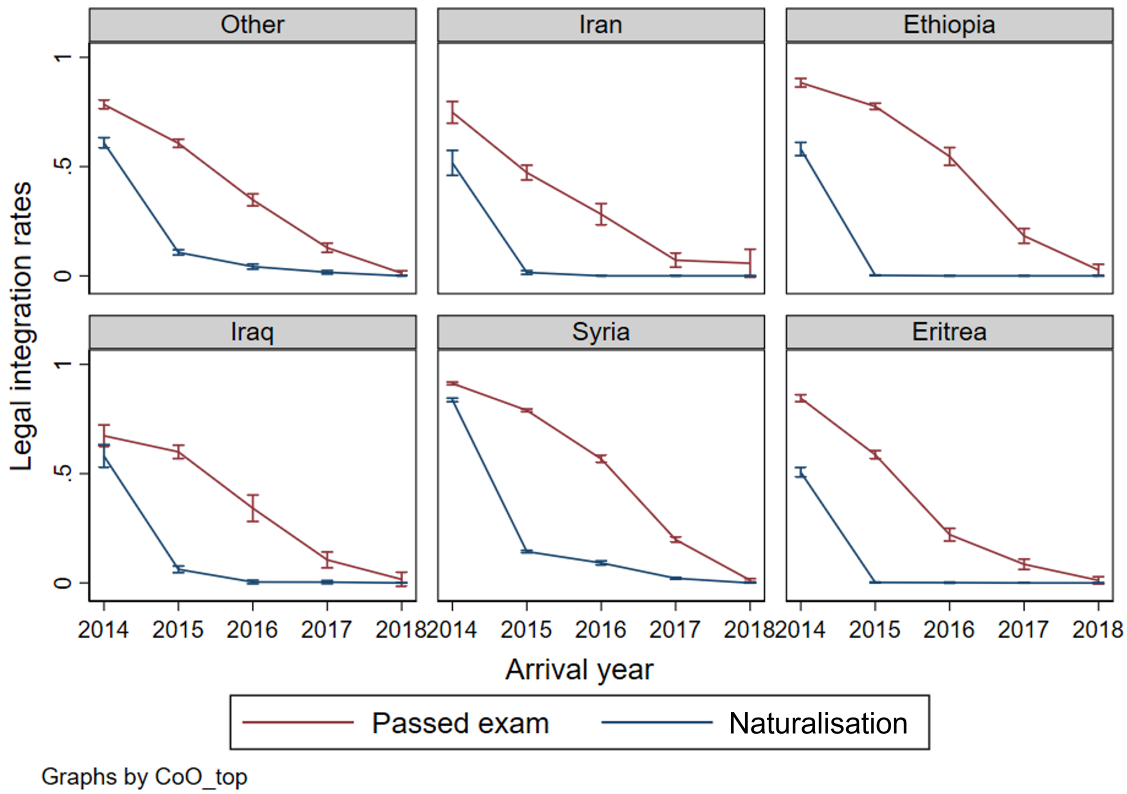
Table A.5: Summary statistics (Urban and non-urban)

	(1) Non-urban		(2) Urban	
	Mean	SD	Mean	SD
Passed exam/Exempted	66.75%		65.15%	
Naturalised	24.18%		23.47%	
Years since migration	4.66	1.07	4.72	1.02
Age at arrival	31.34	11.43	30.29	11.73
Female	36.40%		31.82%	
Education in NL				
No education	81.29%		78.15%	
Low or middle	17.58%		19.16%	
High	1.12%		2.69%	
Country of origin				
Syria	62.32%		58.72%	
Ethiopia	12.30%		9.71%	
Eritrea	9.17%		10.39%	
Iraq	3.35%		3.52%	
Iran	2.29%		3.85%	
Economic conditions				
Duration of social benefits	66.61%	25.92%	64.67%	26.00%
Currently employed	23.25%		27.17%	
N	19,823		28,178	

Note: Column (1) presents summary statistics of those who are allocated into non-urban areas, and Column (2) shows those allocated to urban areas. Urban is defined based on population density: urban (1500 and more addresses /km2, i.e., very strong and strong urban of the COA classification); non-urban (less than 1500 addresses/km2, i.e., moderate, weak and non-urban of the classification).

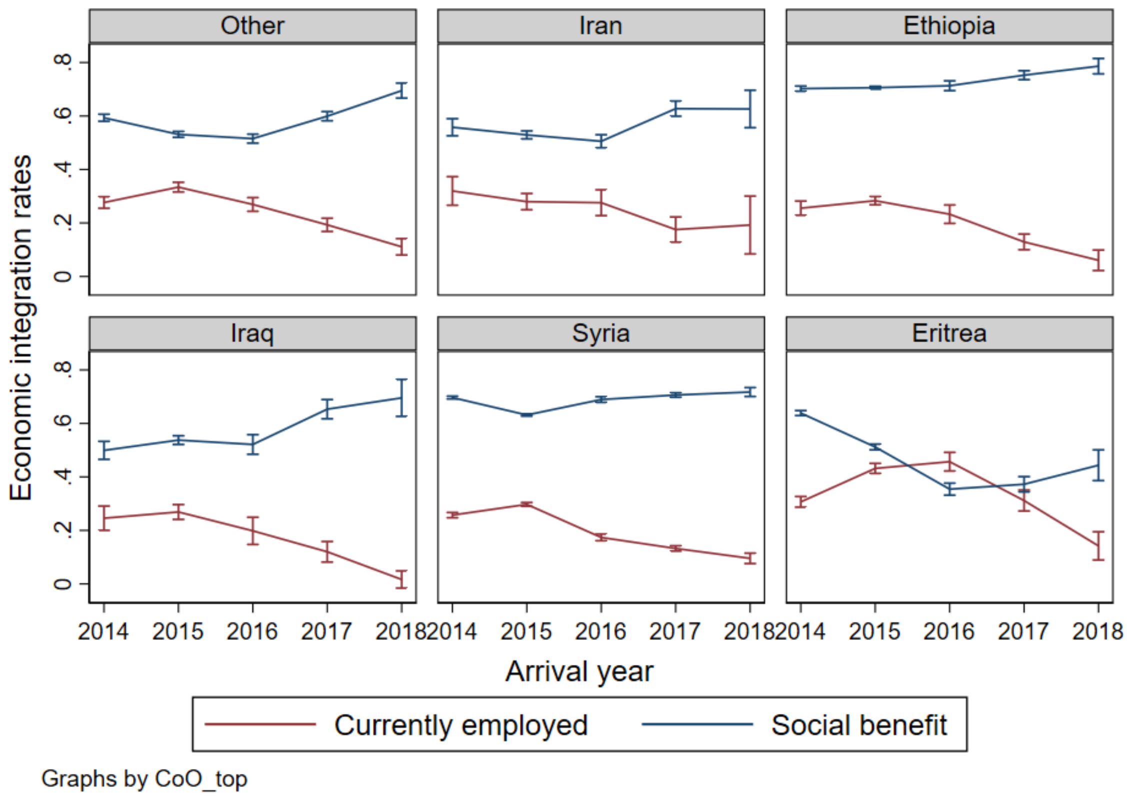
A.3 Integration rates and asylum application periods by country of origin

Figure A.1: Trends of integration rates by country of origin



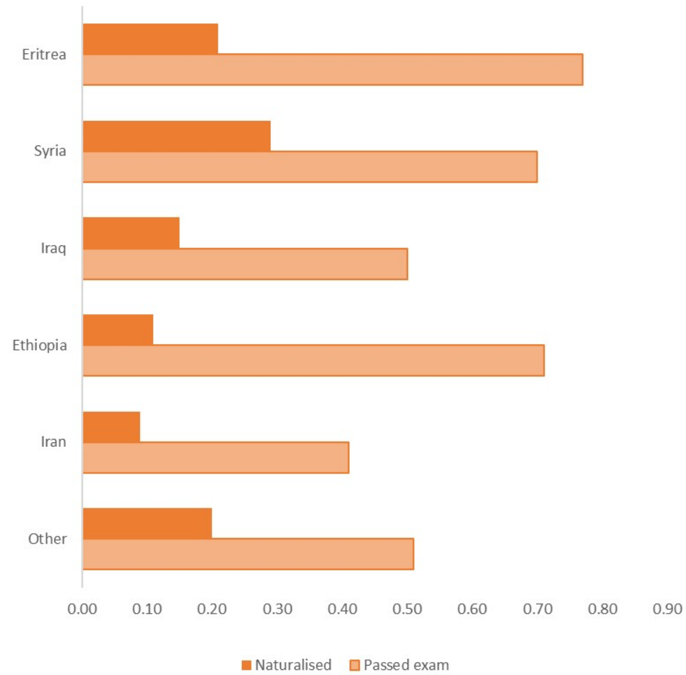
Note: Iran, Ethiopia, Iraq, Syria and Eritrea are the five largest countries of origin among the total sample. 'Other' refers to all countries other than these five.

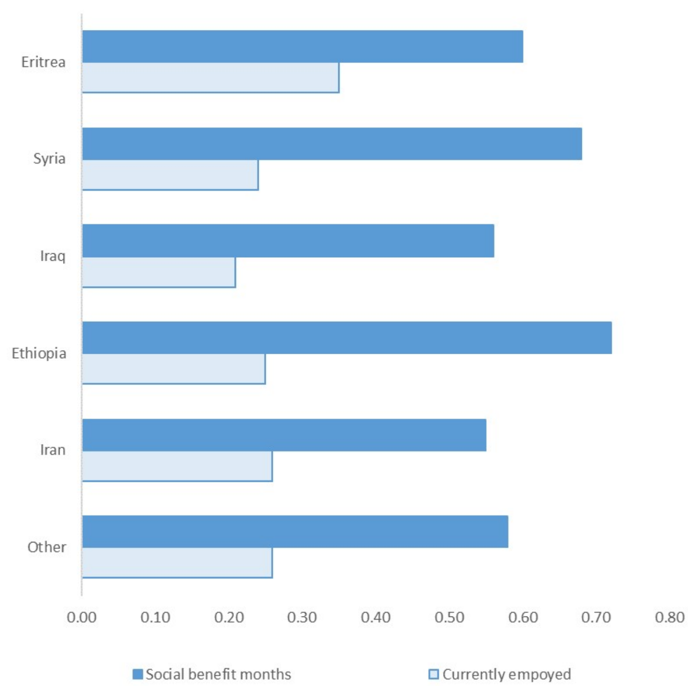
Figure A.2: Trends of economic conditions by country of origin



Note: Iran, Ethiopia, Iraq, Syria and Eritrea are the five largest countries of origin among the total sample. 'Other' refers to countries other than these five.

Figure A.3: Legal and economic indicators by country of origin (Cohort 2014-2020)





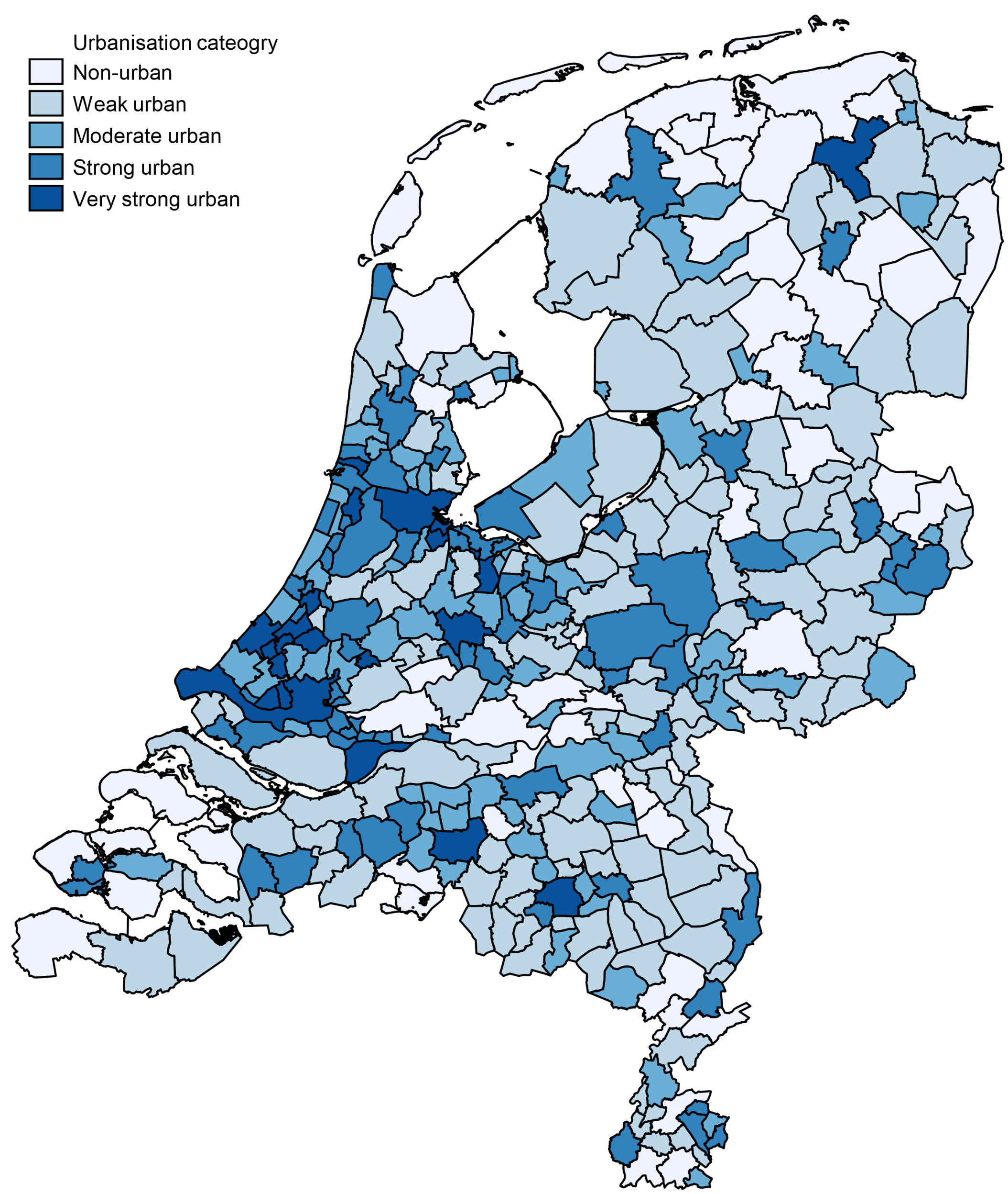
Note: Panel (a) presents legal integration rates by country of origin and Panel (b) exhibits social benefit months provides the number of months that individuals received social benefits as a share of their time in the Netherlands of Cohort 2014-2020. The top five largest refugee nationalities are shown in this figure while other countries are shown as Other.

A.4 Locations of COA centres and sample distribution

COA centres, of which there are 60 as of 2020, are dispersed across the Netherlands. However, most of them are located in less urbanised municipalities. Each COA centre's details are mentioned below and the locations with urbanisation categories on Figure A.5 shows each COA centre's location.

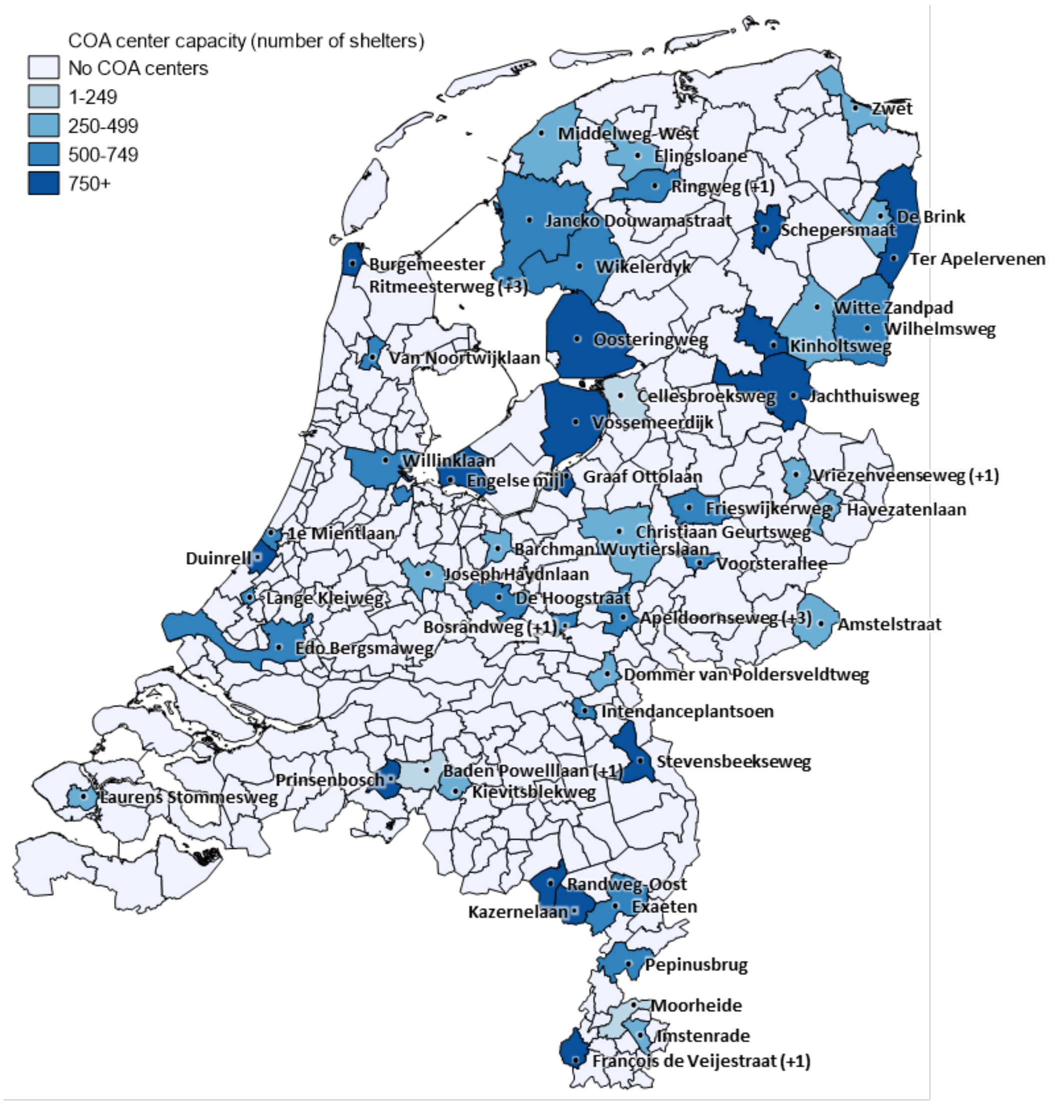
Also, the distribution of our sample, and the Dutch population, by urbanisation categories are shown in Table A.6.

Figure A.4: Urbanisation category by municipalities



Note: It presents urbanisation (population density) by municipalities in 2019 while Panel (b) shows locations and capacities of COA centres as of 2020. The map information comes from Land Registry/Statistics Netherlands, 2020. Panel (a): The blue colour shows urbanisation categories defined by COA based on population density: Very strong urban (2500 and more addresses $/km^2$); Strong urban (1500-2500 addresses $/km^2$); Moderate urban (1000-1500 addresses $/km^2$); Weak urban (500-1000 addresses $/km^2$); and Non-urban (less than 500 addresses $/km^2$). The information comes from Wijk- en buurtkaart 2019 (https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/wijk-en-buurtkaart-2019).

Figure A.5: COA centres by municipalities



Note: The blue colour refers to the capacity (i.e. number of shelters) that the municipalities have as of 2020. The text on the map provides the names of each COA centre(s). The dots in a municipality with a COA present are not intended to reflect the precise locations of each COA centre within that municipality.

Table A.6: Distribution of municipalities and population by urbanisation categories

	Number of Municipalities	Population in the NL (Total)	Population in the NL (Aged 20+)	Sample (Current)
	(1)	(2)	(3)	(4)
Very strong urban	21	25.1%	25.3%	27.8%
Strong urban	74	30.5%	30.4%	30.09%
Moderate urban	76	15.4%	15.3%	14.8%
Weak urban	128	21.3%	21.3%	19.9%
Non-urban	56	7.7%	7.7%	6.5%
Total	355	100.0%	100.0%	100.0%

Note: According to CBS, urbanisation categories are defined into five groups: Very strong (2,500 and over/km^2); Strong (1500 to 2500/km^2); Moderate (1000 to 1500/km^2); Weak (500 to 1000/km^2); and Non-urban (less than 500/km^2). Total population and Population aged 20 and over (i.e. Columns (2) and (3)) are extracted from StatLine as of 2020. As our sample is aged 18 and over, the corresponding population in the Netherlands is a population aged 20 and over. Due to the data limitation, there is no population data based on more detailed age groups. The sample information (i.e. Column (4)) comes from Asielcohort. 'Current' refers to municipalities that the individual lives when passed the integration exam.

A.5 Full regression tables

Due to space limitations in the main text, some variables of the regression tables were omitted and referred to as additional controls. Here we provide the full regression table associated with Tables 5.

Table A.7: Baseline regressions - individual characteristics (*Passed Exam*)

Dep Var: Passed Exam	(1)	(2)	AME
Asylum application months	-0.0570***	-0.0372***	-0.00763
	(0.00158)	(0.00173)	(0.000363)
Years since migration	0.844***	0.872***	0.179
	(0.0387)	(0.0415)	(0.00804)
Arrival age	0.209***	0.139***	0.0284
	(0.00560)	(0.00483)	(0.000992)
Squared arrival age	-0.00275***	-0.00179***	-0.000368
	(7.20e-05)	(6.24e-05)	(1.28e-05)
Female	-0.0553***	-0.0461***	-0.00950
	(0.0167)	(0.0174)	(0.00359)
Education in NL (ref: No education)			
Low/Middle	0.129***	0.385***	0.0746
	(0.0428)	(0.0348)	(0.00637)
High	0.475***	0.714***	0.129
	(0.0577)	(0.0598)	(0.00931)
Duration of social benefits		1.926***	0.395
		(0.0677)	(0.0121)
Currently employed		0.405***	0.0805
		(0.0259)	(0.0510)
Mental health problems		-0.348***	-0.0754
		(0.0339)	(0.00775)
Chronic physical problems		-0.373***	-0.0816
		(0.0635)	(0.0148)
Other controls			
Family size	-0.0231***	0.00477	0.000979
	(0.00555)	(0.00545)	(0.00112)
Position in household (ref: Single)			
Spouse/Partner	0.134***	0.141***	0.0282
	(0.0304)	(0.0295)	(0.00580)
Parent	0.0494**	-0.0167	-0.00346
	(0.0212)	(0.0211)	(0.00436)
Child	-0.470***	-0.154***	-0.0327
	(0.0455)	(0.0433)	(0.00950)
Arrival year (ref: 2014)			
2015	0.387***	0.402***	0.0897

(Continued Table A.7)

	(1)	(2)	AME
	(0.0476)	(0.0521)	(0.0110)
2016	0.285***	0.294***	0.0669
	(0.0834)	(0.0891)	(0.0193)
2017	-0.251**	-0.256**	-0.0620
	(0.110)	(0.116)	(0.0291)
2018	-0.771***	-0.741***	-0.184
	(0.175)	(0.181)	(0.0476)
Country of origin (ref: Other)			
Iran	-0.216***	-0.160***	-0.0351
	(0.0446)	(0.0461)	(0.0103)
Ethiopia	-0.0361	-0.155***	0.0340
	(0.0371)	(0.0384)	(0.00848)
Iraq	0.0533	0.0731	0.0154
	(0.0452)	(0.0449)	(0.00936)
Syria	0.324***	0.239***	0.0489
	(0.0268)	(0.0274)	(0.00573)
Eritrea	0.0694	-0.114**	-0.0248
	(0.0448)	(0.0447)	(0.00985)
Stateless	0.405***	0.430***	0.0824
	(0.0302)	(0.0310)	(0.00553)
Constant	-7.200***	-7.594***	
	(0.253)	(0.273)	
Municipality FE	Yes	Yes	
Observations	48,001	48,001	
Pseud R-sq	0.390	0.428	

Note: Robust standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.1$. AME (Average marginal effect) is calculated based on Column (2). The reference groups for Education in the Netherlands, Arrival Year and Countries of Origin are no education, 2014, and Other countries', respectively.

Table A.8: Baseline regressions - individual characteristics (*Naturalisation*)

Dep Var: Naturalised	(1)	(2)	(3)	(4)	(5)	AME
Arrival age	0.0830***	0.0824***	0.0863***	0.0876***	0.0965***	0.0218
	(0.0113)	(0.0116)	(0.0127)	(0.0127)	(0.0131)	(0.00291)
Arrival age^2	-0.000979***	-0.000970***	-0.00101***	-0.00102***	-0.00109***	-0.000247
	(0.000152)	(0.000155)	(0.000170)	(0.000170)	(0.000176)	(3.94e-05)
Female	0.0724*	0.0693*	0.0684	0.0793*	0.182***	0.0411
	(0.0403)	(0.0406)	(0.0447)	(0.0456)	(0.0464)	(0.0104)
Country of origin (ref: Other)						
Iran	-0.00856	0.0224	0.0484	0.106	0.0930	0.0210
	(0.104)	(0.103)	(0.114)	(0.114)	(0.111)	(0.0250)
Ethiopia	-0.132*	-0.174***	-0.207***	-0.247***	-0.230***	-0.0518
	(0.0693)	(0.0658)	(0.0694)	(0.0705)	(0.0716)	(0.0162)
Iraq	0.233*	0.321**	0.281**	0.305**	0.297**	0.0670
	(0.129)	(0.137)	(0.140)	(0.143)	(0.141)	(0.0317)
Syria	0.730***	0.658***	0.683***	0.683***	0.708***	0.160
	(0.0611)	(0.0592)	(0.0588)	(0.0592)	(0.0600)	(0.0129)
Eritrea	-0.171***	-0.214***	-0.270***	-0.300***	-0.276***	-0.0623
	(0.0587)	(0.0563)	(0.0630)	(0.0623)	(0.0632)	(0.0142)
Education in NL (ref: No education)						
Low/Middle	0.0516	0.0504	0.0659	0.0622	-0.110*	-0.0247
	(0.0503)	(0.0502)	(0.0566)	(0.0566)	(0.0569)	(0.0129)
High	0.563***	0.566***	0.599***	0.598***	0.486***	0.110
	(0.105)	(0.108)	(0.114)	(0.116)	(0.120)	(0.0270)
Years since migration	0.845***	0.839***	0.925***	0.954***	0.903***	0.204
	(0.111)	(0.113)	(0.118)	(0.116)	(0.116)	(0.0256)
Asylum application months		-0.0241***	-0.0247***	-0.0259***	-0.0331***	-0.00747
		(0.00564)	(0.00645)	(0.00663)	(0.00665)	(0.00148)
Mental health problems				-0.464***	-0.396***	-0.0894
				(0.0610)	(0.0643)	(0.0143)
Chronic physical problems				0.0173	0.00385	0.000868
				(0.190)	(0.197)	(0.0445)

(Continued Table A.8)

	(1)	(2)	(3)	(4)	(5)	AME
Duration of social benefits					-1.385***	-0.313
					(0.152)	(0.0337)
Currently employed					0.165***	0.0372
					(0.0415)	(0.00928)
Other controls						
Family size	-0.0219*	-0.0220*	-0.0275*	-0.0242*	-0.0258*	-0.00583
	(0.0132)	(0.0132)	(0.0142)	(0.0146)	(0.0150)	(0.00339)
Position in household (ref: Single)						
Spouse/Partner	0.0641	0.0642	0.126	0.111	0.0874	0.0197
	(0.0724)	(0.0719)	(0.0774)	(0.0785)	(0.0812)	(0.0183)
Parent	0.137***	0.128***	0.137**	0.117**	0.172***	0.0387
	(0.0489)	(0.0489)	(0.0534)	(0.0543)	(0.0526)	(0.0119)
Child	0.0209	-0.000750	-0.0264	-0.0480	-0.224*	-0.0507
	(0.125)	(0.126)	(0.130)	(0.130)	(0.133)	(0.0300)
Stateless	1.027***	1.029***	1.067***	1.092***	1.065***	0.240
	(0.0815)	(0.0823)	(0.0885)	(0.0882)	(0.0899)	(0.0198)
Municipality FE	No	No	Yes	Yes	Yes	
Constant	-6.163***	-5.967***	-7.251***	-7.420***	-6.774***	
	(0.729)	(0.754)	(0.838)	(0.838)	(0.857)	
Observations	9,482	9,482	9,060	9,060	9,060	
Pseud R-sq	0.138	0.140	0.178	0.184	0.198	

Robust standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.1$. The sample is fully eligible individuals for naturalisation, who arrived and obtained Asylum Residence permit in 2014. As the sample is limited to those who entered in 2014, the arrival year dummy is omitted. AME (Average marginal effect) is calculated based on Column (5), in which statistical significance presents the same as AME. The reference groups of Education in the Netherlands, and Countries of origin are no education and 'other countries', respectively. Duration of social benefits is calculated based on the entire period of residence in the Netherlands (number of months to receive social benefits)/(number of total months since migration).

Table A.9: Baseline regressions - municipal level (*Naturalisation*)

Dep Var: Naturalised	(1)	(2)	AME
Asylum application months	-0.0246***	-0.0331***	-0.00746
	(0.00645)	(0.00665)	(0.00148)
Years since migration	0.926***	0.904***	0.204
	(0.118)	(0.116)	(0.0256)
Arrival age	0.0861***	0.0964***	0.0217
	(0.0127)	(0.0131)	(0.0291)
Arrival age^2	-0.00100***	-0.00109***	-0.000246
	(0.000170)	(0.000176)	(3.94e-05)
Female	0.0675	0.181***	0.0394
	(0.0446)	(0.0463)	(0.00967)
Education in NL (ref: No education)			
Low/Middle	0.0683	-0.107*	-0.0251
	(0.0567)	(0.0569)	(0.0137)
High	0.599***	0.485***	0.0917
	(0.114)	(0.120)	(0.0185)
Duration of social benefits		-1.387***	-0.0996
		(0.152)	(0.0337)
Currently employed		0.165***	0.0369
		(0.0416)	(0.00909)
Mental health problems		-0.399***	-0.0996
		(0.0644)	(0.0173)
Chronic physical problems		0.00106	0.00238
		(0.0416)	(0.00909)
Municipality FE	Yes	Yes	
Other controls	Yes	Yes	
Observations	9,060	9,060	
Pseudo R-sq	0.179	0.198	

Note: Robust standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.1$. AME (average marginal effect) is calculated based on Column (2). The sample consists of fully eligible individuals for naturalisation, who arrived and obtained Asylum Residence permit in 2014, and passed the integration exam. As the sample is limited to those who entered in 2014, the arrival year dummy is omitted. The reference groups for Education in the Netherlands, Arrival Year and Country of Origin are no education, 2014, and other countries, respectively. Other controls include family size, position in the household, stateless. Duration of social benefits is calculated based on the entire period of residence in the Netherlands (number of months to receive social benefits)/(number of total months since migration).

Table A.10: AME of baseline regressions - municipal level (*Passed Exam*)

Dep Var: Passed Exam	(1) AME	(2) AME	(3) AME
Panel A: Density Channel			
Population density	0.0222***	0.0479***	
	(0.00612)	(0.0124)	
Population density^2		-0.00433***	
		(0.00136)	
Hours worked in urban			0.0133***
			(0.00377)
Hours worked in rural			0.0129***
			(0.00460)
Observations	48,001	48,001	48,001
Psuedo R-sq	0.430	0.431	0.832
Panel B: Co-national Channel			
Co-national density	0.259***	0.751***	0.439***
	(0.0708)	(0.126)	(0.104)
Co-national density^2		-1.468***	-0.765**
		(0.283)	(0.309)
Population density			0.0403***
			(0.0128)
Population density^2			-0.00406***
			(0.00147)
Observations	48,001	48,001	48,001
Psuedo R-sq	0.429	0.430	0.432

Note: Robust standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.1$.

Table A.11: AME of baseline regressions - municipal level (*Naturalisation*)

Dep Var: Naturalised	(1) AME	(2) AME	(3) AME
Panel A: Density Channel			
Population density	0.00234	0.0335***	
	(0.00517)	(0.0113)	
Population density^2		-0.00509***	
		(0.00153)	
Hours worked in urban			0.00694**
			(0.00325)
Hours worked in rural			0.00556
			(0.00378)
Observations	9,482	9,482	9,482
Pseudo R-sq	0.171	0.172	0.171
Panel B: Co-national Channel			
Co-national density	0.0167	0.455***	0.351
	(0.0767)	(0.205)	(0.219)
Co-national density^2		-1.266**	-0.913
		(0.548)	(0.569)
Population density			0.0274**
			(0.0119)
Population density^2			-0.00439***
			(0.00161)
Observations	9,482	9,482	9,482
Pseudo R-sq	0.171	0.171	0.172

Note: Robust standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.1$. AME is average marginal effect.

Table A.12: AME of baseline regressions - municipal level
(*Sub-sample*)

	Passed Exam		Naturalisation			
	(1) No&Low AME	(2) High AME	(3) No&Low AME	(4) High AME	(5) No&Low AME	(6) High AME
Population density	0.0398***	0.0766*			0.0286**	-0.0623
	(0.126)	(0.0406)			(0.0123)	(0.0982)
Population density^2	-0.00393***	-0.0123**			-0.00466***	0.00946
	(0.00146)	(0.00563)			(0.00166)	(0.0132)
Co-national density	0.426***	0.664*			0.316	1.326
	(0.106)	(0.371)			(0.229)	(2.433)
Co-national density^2	-0.734**	-1.050			-0.878	3.766
	(0.313)	(0.900)			(0.605)	(10.87)
Hours worked in urban			0.00611*	0.0424		
			(0.00342)	(0.0268)		
Hours worked in rural			0.00549	-0.0192		
			(0.00391)	(0.0344)		
Observations	47,019	957	9,269	134	9,269	134
Pseudo R-sq	0.432	0.481	0.170	0.415	0.171	0.488

Note: Robust standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.1$. AME is average marginal effect.

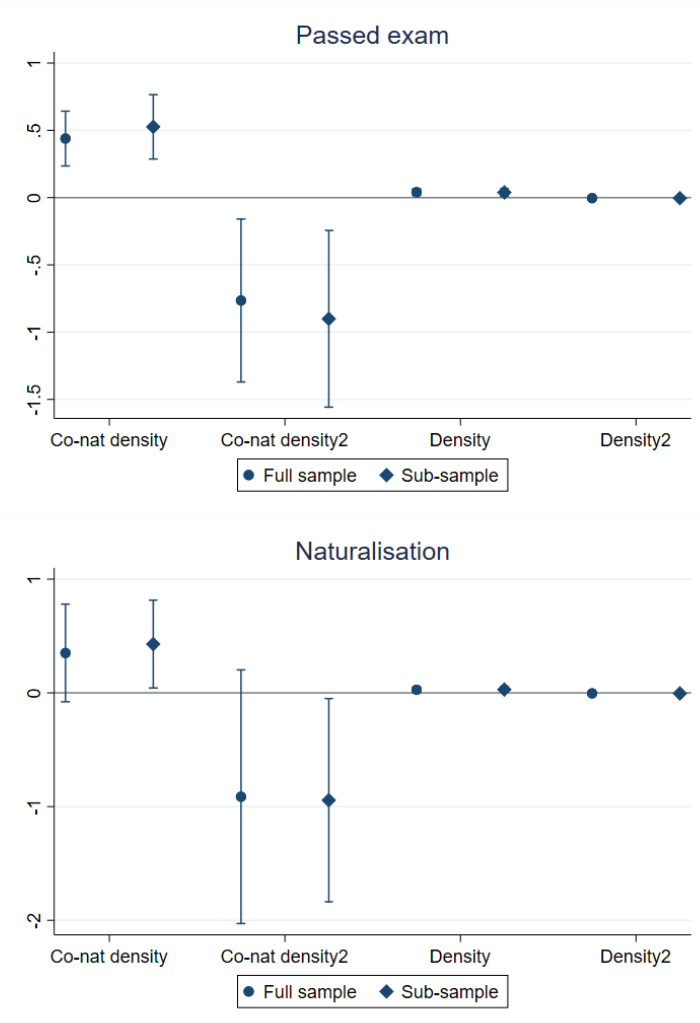
A.6 Robustness check

Table A.13: Robustness check: IV probit

	(1)	(2)	(3)
Panel A: Passed exam			
Population density	0.146***	0.442***	0.440***
	(0.0275)	(0.138)	(0.149)
Population density^2		-0.0393**	-0.0403**
		(0.0172)	(0.0187)
Co-national density			0.122
			(0.843)
Co-national density^2			0.478
			(2.588)
Observations	47,052	47,052	47,052
Wald chi^2	17,126	19,256	19,665
Panel B: Naturalisation			
Population density	-0.0576*	0.0432	0.00232
	(0.0311)	(0.0851)	(0.0921)
Population density^2		-0.0122	-0.0828
		(0.0106)	(0.0116)
Co-national density			2.327**
			(1.017)
Co-national density^2			-5.742**
			(2.625)
Observations	9,290	9,290	9,290
Wald chi^2	1,657	1,657	1,878

Note: All controls and COROP fixed effects are included. The instrumented variables are log(population density) and log(population density)^2 while instrumental variables are log(population density in 1902) and log(population density in 1902)^2.

Figure A.6: AME of base sample and alternative sub-sample



Note: Where dots in the charts show point estimators of average marginal effects, and the lines with the dots mean 95 percent confidence intervals, calculated with standard errors. The base samples are Cohort 2014-2020, aged 18 and over for *Passed Exam* and Cohort 2014, aged 18 and over for *Naturalisation*. The sample sizes are 48,001 and 9,482, respectively. Alternative sub-sample refers to the sample excluding those who are exempt from the integration exam. The sample sizes are 38,001 and 6,569, respectively. We replicate the baseline regressions shown in Column (3) of Table 1.5 and 1.6 with the full set of controls.

Appendix B

Appendix for Chapter 2

B.1 Comparison of integration requirements

Table B.1: Comparison of integration requirements between *Wi*2007 and *Wi*2013

	*Wi*2007	*Wi*2013
Integration period	3.5 years	3 years
Responsibility to find and finance integration courses	Initially own responsibility, with the exception of asylum seekers and ministers of religion. Mandatory offer by municipality for asylum migrants and spiritual leaders, financed by the state.	Own responsibility with loan facility.
Penalties against failure	Financial (administrative fine of up to EUR 1,000 which can be imposed repeatedly) and residence law (no permanent residence permit, naturalisation not possible). (not granting a permanent residence permit, naturalisation not possible without having passed the examination).	Financial (administrative fine of up to EUR 1,250, which can be imposed repeatedly) and residence law (failure to grant a permanent permit for an indefinite period of time or revocation of a temporary permit, naturalisation is not possible without having passed an exam).
Executor	Municipalities	Education Implementation Department *Dienst Uitvoering Onderwijs*: DUO
Exam components	Language (Minimum A2: basic) 1. Test spoken Dutch	Language (Minimum A2: basic) 1. Speaking 2. Reading 3. Listening 4. Writing
	2. Knowledge of Dutch Society (KNS)	5. Knowledge of Dutch Society (KNS): multiple-choice questions within eight topics 1) Work and Income; 2)Values and Norms; 3) Housing; 4) Health and Healthcare; 5) History and Geography; 6) Agencies; 7) Constitution and the rule of law; and 8) Education and upbringing.)
	3. Electronic practical exam based on 4 substantive knowledge profiles	6. Orientation in the Dutch labour market (ONA) (from 2015) 1) Portfolio; 2) 64-hours course or final interview

Note: Source: Table 1 and Table 13 of Algemene Rekenkamer (2017), *Inburgering: Eerste resultaten van de Wet inburgering 2013*. Integration period refers to the given period to pass the exam, which is from the issued date of the resident permit to the date to pass the integration exam.

B.2 Propensity score matching

To mitigate potential bias caused by the over-representation of females and Syrians in the treated group, we select a control group of refugees with similar characteristics to the treated group by using a propensity score matching method. The propensity score is based on the probability of being 'treated' and is estimated as follows:

$$Pr(ONA = 1)_i = \beta IntegPer_i + \gamma X_i' + e_i \tag{B.1}$$

Where ONA is a dummy variable that takes the value of 1 if the individual i is required to sit for the ONA component and 0 otherwise. $Integ\,Per_i$ presents the integration period excluding the extended period due to 'Special circumstances' for individual i[1], and X' is a vector of individual characteristics, including arrival age, gender and country of origin.[2] We perform the matching on the basis of a Gaussian Kernel approach with a bandwidth of 0.01 and common support. This matching approach selects individuals from the control group who are similar to the treated and assigns weights to each control on the basis of the distance in terms of the propensity score. Kernel matching maximises precision by retaining all

[1] The integration period is the duration that the individual took to pass the integration exam since obtaining the residence permit, excluding the extended period due to 'special circumstances' such as late notification of the integration exam. The integration period is used as a proxy for the individual's ability and motivation, which could address the lack of educational attainment outside of the Netherlands.

[2] Caliendo and Kopeinig (2008) suggests that only variables that are not affected by the treatment should be included in the propensity score matching. In other words, variables should either be time-invariant (e.g., arrival age, country of origin) or measured before the treatment. As shown earlier, the integration period without the extended period due to 'special circumstances' does not differ before and after the ONA introduction while months since migration is affected by the permit year because Cohort 2015 extended their integration period due to the 'special circumstances'. Hence, we exclude months since migration from the matching equation.

matched controls and assigning greater weights to better matches (Garrido et al., 2014).[3].

Table B.2 presents the results of balance tests for the applied propensity score matching. As the table suggests, none of the variables show a statistical difference between the matched treatment and control groups for each variable (i.e. P-value of Row (M) shows higher than 0.1 and/or % bias presents less than 5) while the unmatched sample (Row (U)) shows statistically significant differences (Caliendo and Kopeinig, 2008).

[3] Since there seems to be little consensus on which matching approach is preferable, we apply different approaches and bandwidths/calipers: Nearest Neighbour matching with caliper; Radius matching with the optimal caliper (i.e. 0.2 of the standard deviation of propensity score); Kernel with a bandwidth of 0.06; and Kernel with a bandwidth of 0.01 with Mahalanobis matching option. However, the Kernel matching with a bandwidth of 0.01 leads to the best outcome in terms of the quality of the matching, according to the balancing test. A strict bandwidth ensures the matching quality (Li, 2013)

Table B.2: Balance test of propensity score matching

Variable	Unmatched Matched	Mean Treated	Mean Control	%bias	%reduction |bias|	T-test t	T-test p > |t|
Integration period	U	29.205	33.897	-71.0		-42.34	0.000
	M	29.205	29.176	0.4	99.4	0.34	0.738
Asylum claim period	U	2.7094	2.9286	-7.5		-4.41	0.000
	M	2.7094	2.7161	-0.2	96.9	-0.14	0.891
Arrival age	U	30.307	28.863	16.6		9.79	0.000
	M	30.306	30.481	-2.0	87.9	-1.19	0.236
Female	U	.39765	.29127	22.5		13.27	0.000
	M	.39765	.39064	1.5	93.4	0.87	0.385
Country of origin (ref: Other)							
Syria	U	.70579	.5637	29.8		17.64	0.000
	M	.70579	.71112	-1.1	96.2	-0.71	0.477
Eritrea	U	.16635	.27122	-25.6		-15.15	0.000
	M	.16635	.14689	4.7	81.4	3.24	0.001
Iran	U	.01692	.02156	-3.4		-2.00	0.046
	M	.01692	.01798	-0.8	77.3	-0.49	0.626
Afghanistan	U	.00901	.01689	-7.0		-4.15	0.000
	M	.00901	.00943	-0.4	94.6	-0.27	0.788
Iraq	U	.01078	.01945	-7.1		-4.22	0.000
	M	.01078	.01149	-0.6	91.8	-0.41	0.683
Family size	U	1.9234	1.9732	-3.1		-1.85	0.065
	M	1.9234	1.9535	-1.9	39.6	-1.16	0.245
Education in NL (ref: No education)							
Low/Middle	U	.15816	.15031	2.2		1.28	0.200
	M	.15816	.15064	2.1	4.2	1.26	0.208
High	U	.04094	.01658	14.6		8.53	0.000
	M	.04094	.03701	2.4	83.9	1.23	0.219
Position in household (ref: Single)							
Spouse/Partner	U	.07478	.0793	-1.7		-1.00	0.317
	M	.07478	.07095	1.4	15.2	0.89	0.372
Parent	U	.58966	.48319	21.5		12.67	0.000
	M	.58966	.59757	-1.6	92.6	-0.98	0.329
Child(18+)	U	.04531	.0288	8.7		5.13	0.000
	M	.04531	.04325	1.1	87.5	0.61	0.544
Other	U	.0086	.00935	-0.8		-0.47	0.639
	M	.0086	.00837	0.2	69.1	0.15	0.878

Note: Integration period refers to the number of months between the asylum permit being granted and the integration exam being passed, excluding the extended period due to 'special circumstances'.

B.3 Summary statistics of sub-sample

Table B.3: Summary statistics of sub-sample (Country of origin)

	Syria				Other country			
	(1) Treatment		(2) Control		(3) Treatment		(4) Control	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Asylum claim period	2.12	2.12	2.18	1.75	4.20	4.80	4.13	3.46
Integration period	29.39	5.05	29.46	5.34	28.20	4.75	28.47	5.12
Arrival age	31.77	8.85	31.73	8.93	27.76	7.17	28.10	7.90
Female	41.79%		42.18%		35.14%		36.08%	
Country of origin								
Syria	100.00%		100.00%		0.00%		0.00%	
Eritrea	0.00%		0.00%		55.90%		51.53%	
Iran	0.00%		0.00%		3.12%		3.47%	
Family size	1.63	1.35	1.38	1.09	1.11	0.57	1.10	0.48
Education in NL								
No education	98.49%		98.36%		97.36%		96.00%	
Low/Middle	1.44%		1.50%		2.25%		2.72%	
High	0.08%		0.14%		0.38%		1.28%	
Labour market performance								
Emp. prob.	0.02	0.12	0.01	0.08	0.02	0.16	0.02	0.14
Hours worked	50.38	49.39	52.02	54.38	71.68	61.00	70.34	44.00
Hourly wages	10.01	2.98	10.97	3.43	10.97	3.43	10.13	1.70
Observations	5,085		3,657		2,086		2,671	

Note: All summary statistics are computed based on the matched sample and time-variant variables are reported for the month that each individual obtained the asylum residence permit (t_0). Hours worked and hourly wages are conditional on being employed, which reduces the sample size.

Table B.4: Summary statistics of sub-sample (Gender)

	Male				Female			
	(1) Treatment		(2) Control		(3) Treatment		(4) Control	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Asylum claim period	3.41	2.95	3.25	2.44	1.68	3.44	1.81	2.68
Integration period	28.81	4.95	28.91	5.17	29.40	5.04	29.30	5.66
Arrival age	30.58	8.76	30.95	9.06	30.62	8.33	30.72	8.60
Country of origin								
Syria	68.65%		70.34%		74.35%		75.29%	
Eritrea	19.55%		17.43%		11.30%		9.26%	
Iran	1.48%		1.31%		2.06%		2.30%	
Family size	1.72	1.43	1.47	1.22	1.12	0.56	1.08	0.39
Education in NL								
No education	97.59%		96.94%		99.02%		98.21%	
Low/Middle	2.16%		2.28%		0.94%		1.69%	
High	0.26%		0.79%		0.03%		0.10%	
Labour market performance								
Emp. prob.	0.03	0.16	0.02	0.12	0.00	0.07	0.00	0.05
Hours worked	55.21	52.28	60.25	47.99	92.15	70.37	52.08	59.70
Hourly wages	9.96	2.32	10.07	2.46	11.36	5.62	11.15	2.70
Observations	4,312		4,463		2,858		1,865	

Note: All summary statistics are computed based on the matched sample and time-variant variables are reported for the month that each individual obtained the asylum residence permit (t_0). Hours worked and hourly wages are conditional on being employed, which reduces the sample size.

B.4 Full-tables of specifications

Table B.5: Baseline specification

	Pr(employment) (1)	Log(hours worked) (2)	Log(hourly wages) (3)
$Period * ONA$ (ref: $Period_0$)			
$Period_1$	0.00432	0.0531	0.104**
	(0.00380)	(0.128)	(0.0424)
$Period_2$	0.00426	0.0515	0.111**
	(0.00499)	(0.166)	(0.501)
$Period_3$	0.00225	0.107	0.114
	(0.00624)	(0.151)	(0.0740)
$Period_4$	0.0254	0.107	0.0971
	(0.00744)	(0.164)	(0.0874)
$Period_5$	-3.80e-05	0.0931	0.0624
	(0.00824)	(0.161)	(0.0872)
$Period_6$	-0.000351	0.100	0.0764
	(0.00939)	(0.169)	(0.0969)
$Period_7$	0.00557	0.141	0.0581
	(0.0105)	(0.172)	(0.0986)
$Period_8$	0.0179	0.144	0.0460
	(0.0112)	(0.176)	(0.959)
$Period_9$	0.0184	0.0923	0.0420
	(0.0120)	(0.174)	(0.0920)
$Period_{10}$	0.0144	0.0585	0.0420
	(0.0124)	(0.172)	(0.0911)
$Period_{11}$	0.00832	0.0445	0.0216
	(0.0132)	(0.166)	(0.0911)
$Period_{12}$	0.00872	0.0445	0.0216
	(0.0135)	(0.166)	(0.0947)
$Period_{13}$	0.0138	0.0272	0.0337
	(0.0134)	(0.166)	(0.0947)
$Period_{14}$	0.0120	0.0358	0.372
	(0.0134)	(0.163)	(0.0957)
$Period_{15}$	0.0216	0.0397	0.0317
	(0.0141)	(0.163)	(0.0959)
$Period_{16}$	0.0328**	0.0558	0.0233
	(0.0141)	(0.165)	(0.0960)
$Period_{17}$	0.0388***	0.0518	0.0414
	(0.0139)	(0.162)	(0.0993)
$Period_{18}$	0.0524***	0.00776	0.0391
	(0.0141)	(0.166)	(0.0973)
$Period_{19}$	0.0519***	0.0896	0.0449

	Pr(employment)	Log(hours worked)	Log(hourly wages)
	(0.0148)	(0.163)	(0.0969)
$Period_{20}$	0.0564***	0.0744	0.0389
	(0.0149)	(0.162)	(0.0974)
$Period_{21}$	0.0616***	0.0854	0.0585
	(0.0144)	(0.163)	(0.0960)
$Period_{22}$	0.0638***	0.106	0.0556
	(0.0140)	(0.165)	(0.0975)
$Period_{23}$	0.0641***	0.124	0.0705
	(0.0161)	(0.167)	(0.0990)
$Period$ (ref: $Period_0$)			
$Period_1$	-0.00831**	-0.0107	-0.0674*
	(0.00389)	(0.108)	(0.0358)
$Period_2$	-0.00602	0.0154	-0.0820*
	(0.00561)	(0.147)	(0.0452)
$Period_3$	-0.00387	0.0150	-0.105
	(0.00648)	(0.136)	(0.0660)
$Period_4$	0.00126	0.0404	-0.0682
	(0.00854)	(0.158)	(0.0797)
$Period_5$	0.0110	0.124	-0.0347
	(0.00971)	(0.159)	(0.0853)
$Period_6$	0.0193*	0.196	-0.0276
	(0.0107)	(0.168)	(0.100)
$Period_7$	0.0319***	0.253	0.00630
	(0.0112)	(0.168)	(0.100)
$Period_8$	0.0436***	0.336**	0.0178
	(0.0121)	(0.166)	(0.106)
$Period_9$	0.0640***	0.425**	0.0325
	(0.0127)	(0.165)	(0.104)
$Period_{10}$	0.0899***	0.519***	0.0508
	(0.0130)	(0.167)	(0.104)
$Period_{11}$	0.115***	0.576***	0.0827
	(0.0131)	(0.164)	(0.107)
$Period_{12}$	0.131***	0.592***	0.102
	(0.0135)	(0.167)	(0.109)
$Period_{13}$	0.135***	0.661***	0.112
	(0.0136)	(0.167)	(0.110)
$Period_{14}$	0.151***	0.703***	0.126
	(0.0138)	(0.168)	(0.110)
$Period_{15}$	0.150***	0.703***	0.148
	(0.0142)	(0.169)	(0.110)
$Period_{16}$	0.144***	0.714***	0.171
	(0.0136)	(0.176)	(0.113)
$Period_{17}$	0.147***	0.749***	0.178
	(0.0136)	(0.171)	(0.114)
$Period_{18}$	0.141***	0.744***	0.201*

	Pr(employment)	Log(hours worked)	Log(hourly wages)
	(0.0138)	(0.176)	(0.113)
$Period_{19}$	0.146***	0.774***	0.208*
	(0.0134)	(0.173)	(0.113)
$Period_{20}$	0.156***	0.804***	0.233**
	(0.0130)	(0.173)	(0.113)
$Period_{21}$	0.164***	0.827***	0.245**
	(0.0127)	(0.175)	(0.113)
$Period_{22}$	0.171***	0.842***	0.261**
	(0.0123)	(0.177)	(0.113)
$Period_{23}$	0.181***	0.849***	0.281**
	(0.0128)	(0.176)	(0.114)
Education in NL (ref: No education)			
Low/Middle	0.162***	0.0742*	-0.0752***
	(0.0101)	(0.0379)	(0.0120)
High	0.0904***	-0.0909	-0.00937
	(0.0137)	(0.0799)	(0.0476)
Health conditions			
Mental health issues	-0.0706***	-0.229***	-0.0427
	(0.0131)	(0.0743)	(0.0433)
Chronic physical issues	-0.0701*	0.156	0.0990
	(0.0370)	(0.142)	(0.241)
Position in household (ref: Single)			
Spouse/Partner	0.00309	0.0222	-0.00611
	(0.00704)	(0.0196)	(0.0105)
Parent	-0.0666***	-0.0418*	-0.0322***
	(0.00745)	(0.0251)	(0.0108)
Child(18+)	-0.0321***	0.00643	-0.0730***
	(0.0159)	(0.0449)	(0.0254)
Other	-0.0257***	0.0133	-0.0262
	(0.00952)	(0.0300)	(0.0163)
Constant	0.183***	3.772***	2.374***
	(0.00832)	(0.111)	(0.0678)
Observations	938,991	257,554	257,554
R-squared	0.515	0.556	0.520
Individual FEs	Yes	Yes	Yes
Municipality FEs	Yes	Yes	Yes
Period FEs	Y-Q-C	Y-Q-C	Y-Q-C

Note: Robust standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.10$. All estimations include individual FEs, municipality FEs and Year-Quarter-Country FEs. Standard errors are clustered at the municipality level. The reference period is t_0 for both the period dummies and the interaction of the period and treatment.

B.5 Employment stability measurement

Herfindahl-Hirschman Index (HHI) is a measure of the level of concentration applied to market dominance or diversity in labour forces (Rhoades, 1993; Ozgen, 2021). In this study, it is used to gauge employment stability, in other words job concentration in the individuals' entire tenure in the Netherlands. The equation is as follows.

$$HHI_{i,t} = \sum_{j=1}^{J} s^2_{j(i,t)} \tag{B.2}$$

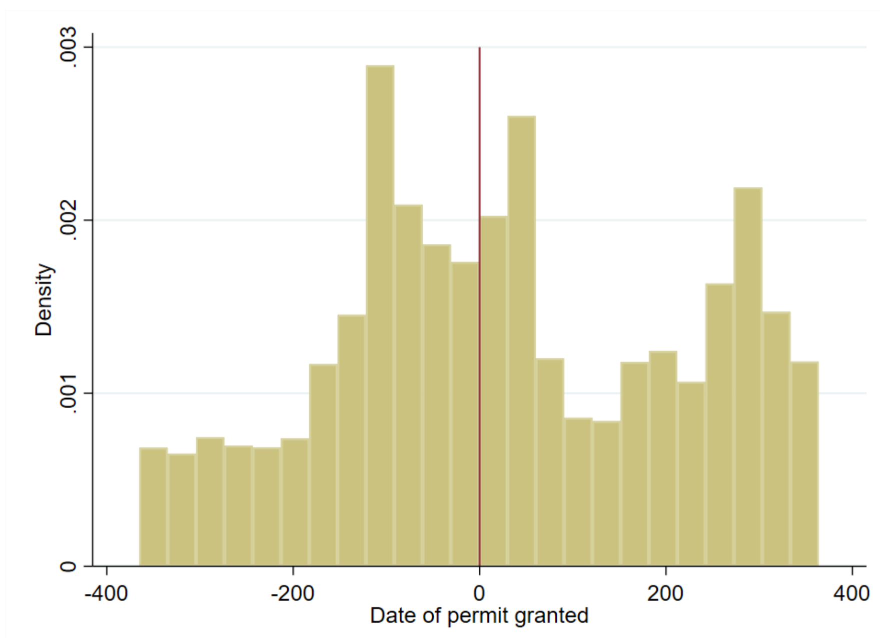
where, $s_{j(i,t)}$ refers to the share of job j of individual i at time t out of the total tenure, and J is a total number of jobs that an individual i has experienced by time t. If an individual has been employed in a single job throughout the period, HHI takes a value of one. In contrast, the more an individual changes jobs the closer the index goes towards zero.

B.6 Robustness checks

B.6.1 Regression Discontinuous Design (RDD)

Before estimating the ONA impact with Regression Discontinuity (RD) estimation, we investigate whether the treatment and control groups around the cutoff point are different in observable characteristics from each other. Figure B.1 presents the distribution of individuals by the issue date of the asylum residence permit while Table B.6 shows the balancing test around the cutoff point at the time t_0, suggesting they are not statistically different in terms of observable characteristics.

Figure B.1: Distribution of the issue date of asylum residence permit



To specify the causal impact of the ONA, we estimate an RD model as specified below.

Table B.6: Balancing tests for RDD

	Mean (1)	SD (2)	RD Estimate (3)	Confidence Interval (4)
Asylum claim period	2.81	2.94	-0.14	[-0.53 ; 0.24]
Arrival age	30.01	8.53	-1.42	[-2.73 ; -0.12]
Female	0.35	0.48	-0.02	[-0.11 ; 0.07]
Country of origin				
Syria	0.65	0.48	-0.03	[-0.09 ; 0.02]
Eritrea	0.21	0.41	0.11	[0.08 ; 0.14]
Iran	0.02	0.14	-0.01	[-0.03 ; 0.01]
Family size	1.36	1.06	-0.03	[-0.19 ; 0.14]
Education in NL				
No education	0.98	0.14	-0.01	[-0.04 ; 0.01]
Low/Middle	0.02	0.13	0.02	[0.00 ; 0.05]

Note: The number of effective observations i.e. within the bandwidth is 4,571 while the total number of observations is 13,501 (individuals at time t_0). The mean squared error (MSE) optimal bandwidth from Calonico et al. (2019) is applied while standard errors are clustered at the municipality level.

$$Y_{it} = \alpha + \delta ONA_i + \eta(x_i - c) + \lambda ONA_i(x_i - c) + \rho Z_{it} + \epsilon_{it} \tag{B.3}$$

Where $x_i - c$ presents the distance in time between the cutoff date and the date the individual was granted the asylum residence permit, and Z is the duration in the Netherlands. Since the treatment and control groups are similar around the cutoff point without matching, it is not necessary to control for individual characteristics.

The coefficient δ is the regression discontinuity estimate and shows the causal intention to treat (ITT) effect of ONA on labour market outcomes. We estimate this specification for the whole period (i.e. $Period_0$ to $Period_{23}$), and also estimate separate regressions in each quarter after the integration period began, using Weighted Least Squares with a triangular kernel to give more weight to the observations closer to the cutoff. The mean squared error (MSE) optimal bandwidth from Calonico et al. (2019) is used while standard errors are clustered at the municipality level as in the baseline estimation to maintain consistency.

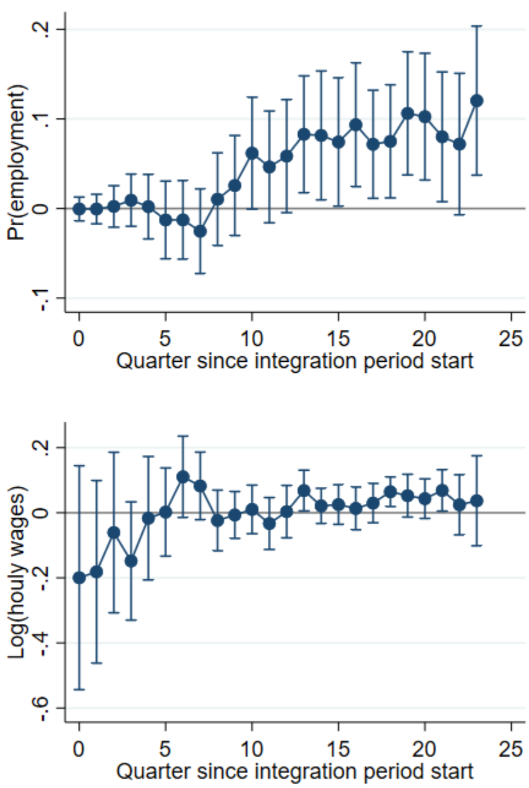
Table B.7 presents the average impacts of the ONA over the period while the ONA impacts per quarter are shown in Figure B.2. Both estimation results support the baseline implications: the ONA boosts employment probability by on average 5.5 percentage points over 8 years, but has no impacts on hours worked or hourly wages (Table B.7). The significantly positive impacts on employment probability appear around 15 quarters after the asylum application (Figure B.3).

Table B.7: Impact of the ONA (whole period)

	Pr(employment) (1)	Log(hours worked) (2)	Log(hourly wages) (3)
RD Estimate	0.0448**	0.0613	0.0187
	(0.0193)	(0.0521)	(0.0198)
Bandwidth	81.42	87.20	89.74
Month since migration	Yes	Yes	Yes
Total observations	939,056	257,723	257,723
Effective Observations	305,899	91,028	91,767

Note:Standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.10$. The effective observations refer to the number of observations within the bandwidth. The mean squared error (MSE) optimal bandwidth from Calonico et al. (2019) is applied while standard errors are clustered at the municipality level. The outcomes (i.e. coefficients of δ from Equation (B.3) shown in RD Estimate row) are the average of outcomes over 8 years (i.e. $Period_0$ to $Period_{23}$).

Figure B.2: RD estimates by quarters between 2014-2021



Notes: The economic outcomes impacts of the ONA (i.e. coefficients of δ from Equation (B.3)) are plotted by quarters.

B.6.2 A Differences-in-Differences Approach

Our findings so far suggest that the ONA component may have had a positive statistically significant impact on employment probability but not on hours worked or hourly wages. To explore this finding further, we focus on the precise timing at which each individual passed the exam and utilise a Differences-in-Differences (DiD) approach.

Using the same matched sample and period of analysis we amend Equation (2.2) to replace the period dummies with a dummy indicating when the individual passed the exam, as outlined in Equation (B.4).

$$Y_{it} = \alpha + \delta Post_{it} + \eta(Post_{it} * ONA_i) + \rho Z_{it} + \mu_i + \tau_t + \sigma_{it} + \epsilon_{it} \tag{B.4}$$

Where $Post$ is a dummy variable taking the value of 1 if the individual passed the exam in that period, otherwise 0. Coefficient η now captures the impact of the ONA on labour market outcomes.[4] Apart from $Post$, other variables are consistent with those included in our baseline estimations although we now include duration since migration within our time-variant individual controls (i.e. Z). In Equation (2.2), the multiple period dummies controlled for the effect of time since migration whereas the binary post-exam dummy in Equation (B.4) is not able to do so, hence the inclusion of this additional control. We also include Year-Month (Y-M) fixed effects.

[4] Due to data availability, we have no information about the exact timings when each individual exam component has been passed. According to DUO, however, refugees tend to take ONA as the last exam component because it requires a certain level of Dutch proficiency. Hence, we can assume the month when the individual passed the exam is the time they completed ONA.

Table B.8: DiD estimation

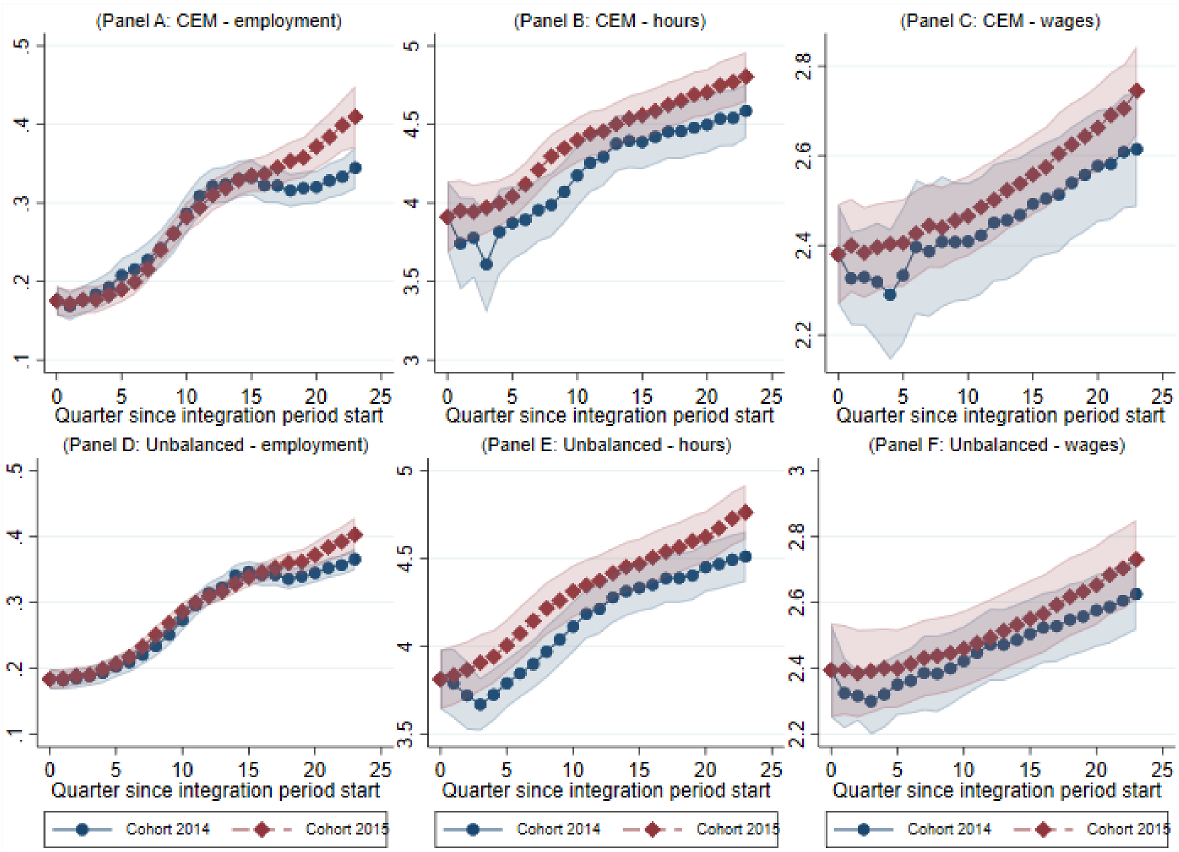
	(1)	(2)	(3)	(4)
Panel A: (Dep. Var.) Pr(employment)				
Post	0.0629***	0.0615***	0.0618***	0.0571***
	(0.00574)	(0.00577)	(0.00576)	(0.00582)
Post*ONA	0.0242***	0.0243***	0.0244***	0.0287***
	(0.00657)	(0.00658)	(0.00657)	(0.00652)
Observation	938,991	938,991	938,991	938,991
Panel B: (Dep. Var.) Log(hours worked)				
Post	0.184***	0.180***	0.179***	0.169***
	(0.0292)	(0.0296)	(0.0294)	(0.0290)
Post*ONA	-0.0192	-0.0188	-0.0184	-0.00492
	(0.0339)	(0.0341)	(0.0340)	(0.0330)
Observation	257,555	257,555	257,555	257,554
Panel C: (Dep. Var.) Log(hourly wage)				
Post	0.0164	0.0144	0.0137	0.0136
	(0.0165)	(0.0166)	(0.0165)	(0.0157)
Post*ONA	-0.0308*	-0.0307*	-0.0303*	-0.0284*
	(0.0173)	(0.0173)	(0.0173)	(0.0170)
Observation	257,555	257,555	257,555	257,554
Month since migration	No	No	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Period FE	Y-Q	Y-M	Y-Q	Y-Q-C
Individual FE	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes

Note: Robust standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.10$. Y-Q and Y-M refers to Year-Quarter and Year-Month fixed effects respectively whereas Y-Q-C refers to Year-Quarter-Country of origin fixed effects. The estimations for log(hours worked) and log(hourly wages) are conditional on being employed, which decreases the sample size.

Table B.8 presents the results of the DiD estimation. *Panel A*, which shows the impacts on employment probabilities, indicates the significant positive effects of the ONA. Across the regression patterns, the interaction coefficients remain in the range of 0.0242 to 0.0287 at the 0.01 significance level. In other words, the ONA boosts the employment probabilities of Cohort 2015 by about 2.42 to 2.87 percentage points after its completion. *Panel B* and *C* provide estimation results for hours worked and hourly wages, suggesting no significant impact of the ONA on these two outcomes across the specifications. From the DiD estimations, we observe that the ONA improves refugees' probability of employment, yet provides no significant impact on the quantity and quality of employment. These implications are consistent with the baseline results from Equation (2.2).

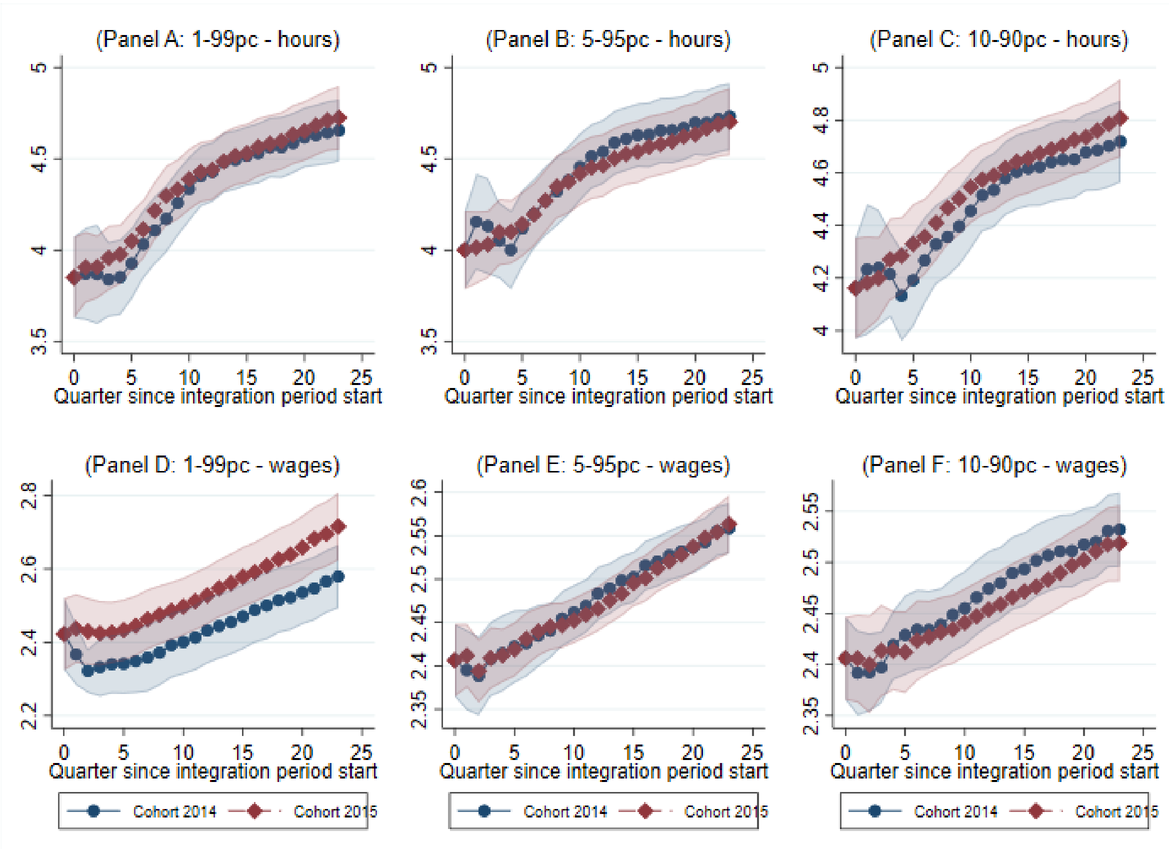
B.6.3 Other robustness checks

Figure B.3: Different panel



Note: Robust standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.10$. All regressions include $Month\, Since\, Migration$, Individual FE, Municipality FE and Year-Quarter-Country FE. Panel A-C show the estimation results based on the matched sample with Coarse Exact Matching (CEM) (Panel A: 617,960 observations, Panel B and C: 170,397 observations), while Panel D-F are the unbalanced sample (Panel D: 1,312,674 observations, Panel E and F: 368,221 observations).

Figure B.4: Excluding outliers



Note: Robust standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.10$. All regressions include $Month\,Since\,Migration$, Individual FE, Municipality FE and Year-Quarter-Country FE. Panel A and D are based on the sample within 1-99 pc (Panel A: 251,904 observations, Panel B: 252,264 observations), Panel B and E are 5-95 pc (Panel C: 231,124 observations, Panel D: 234,973 observations), and the remaining panels presents the results of 10-90 pc (Panel E: 203,881 observations, Panel F: 210,646 observations).

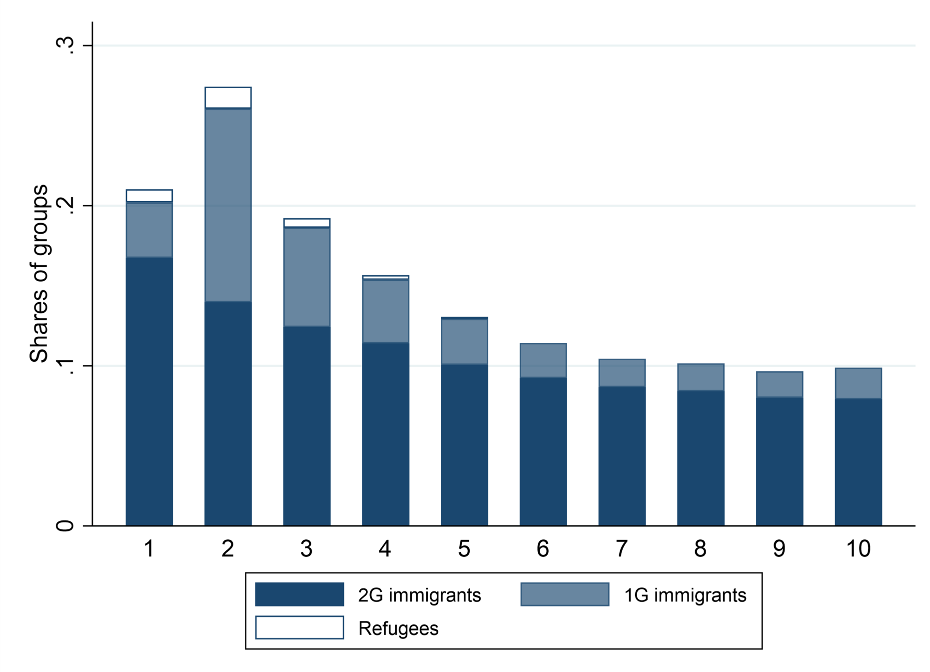
Appendix C

Appendix for Chapter 3

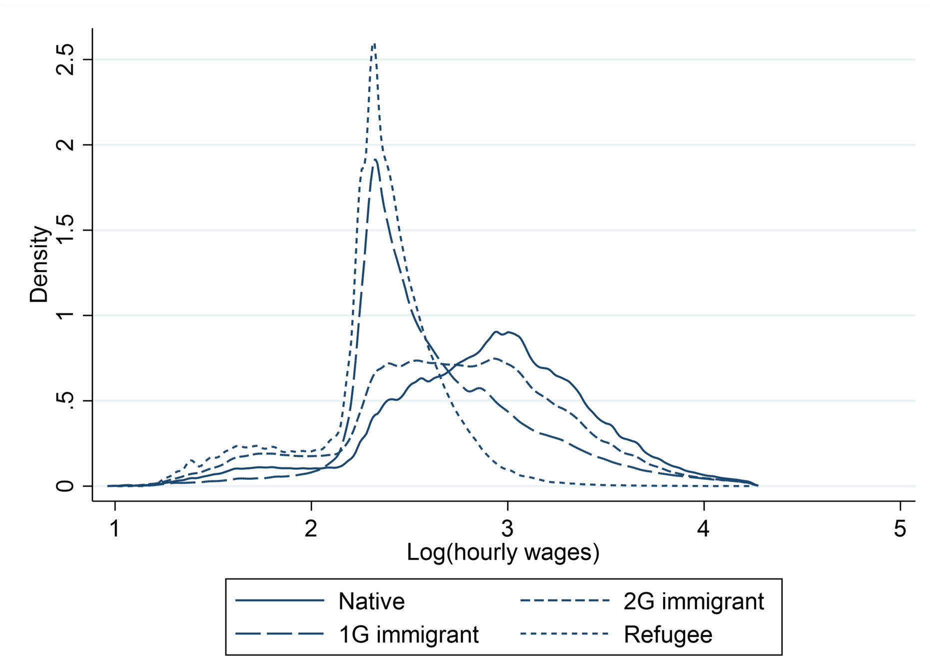
C.0.1 Descriptive analysis

Figure C.1: Wage distribution by wage decile and by group

(a) Group composition by wage decile

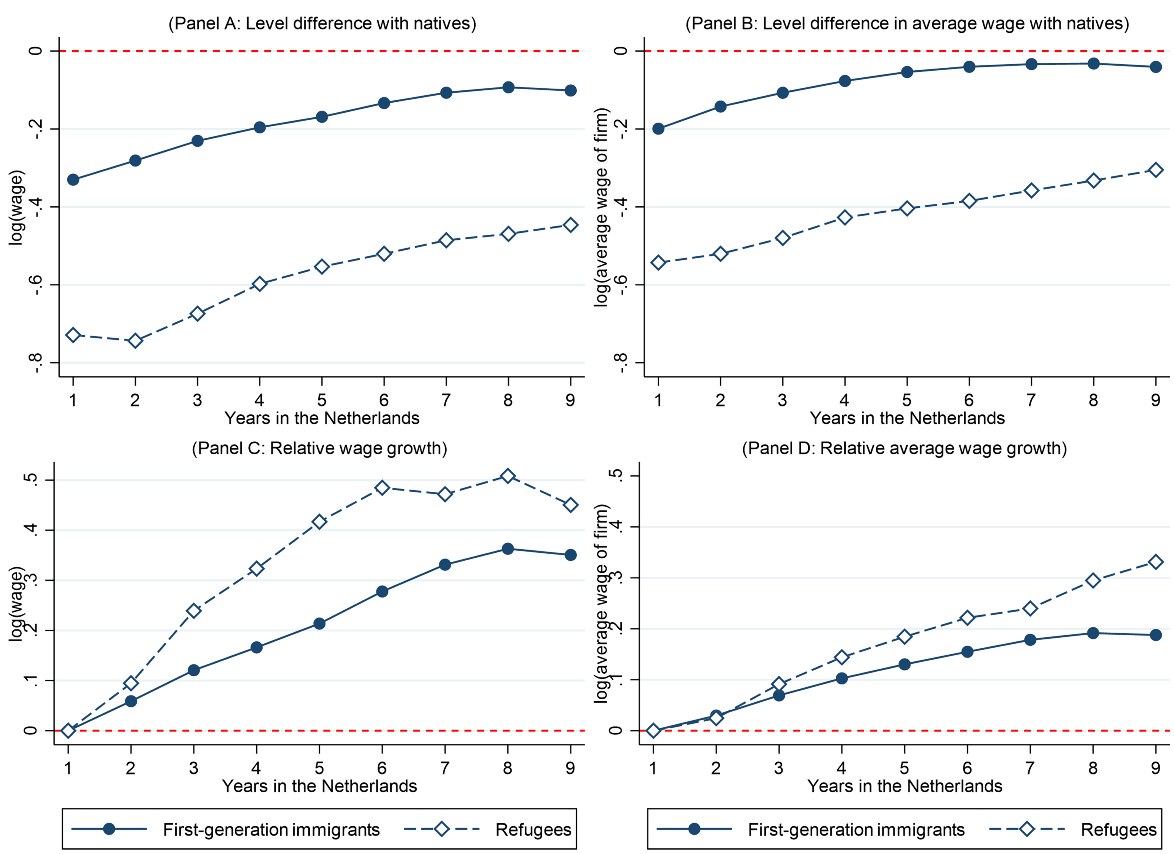


(b) Wage distribution among sample by group



Note: The bars of Panel A show the composition of each decile in the hourly wage distribution in terms of immigrant groups. As natives are dominant in all deciles, native compositions are hidden from the graph. The lines of Panel B present sample distributions by groups.

Figure C.2: Wage gap and years in the Netherlands



Note: Panel A shows wage gaps between natives and first-generation immigrants or refugees over the years since migration. The wage gaps are calculated as the wages of the individual i at year t minus the average wages of natives in the same year. Panel B reports gaps in average wages of the firms in which each individual i works. Panel C exhibits relative wage growth, setting the wages in the first year in the Netherlands as the reference (i.e. $log(w_{it}) - log(w_{i1})$), while Panel D presents the growth of average wages of firms in which each individual i works.

Table C.1: Summary statistics of dual-connected set

	Dual-connected set with financial data					
	Native & 2nd-gen		Native & 1st-gen		Native & Refugee	
	Native	Immig.	Native	Immig.	Native	Immig.
Individual level						
Age	38.33	31.66	37.80	31.26	36.82	27.41
(std.dev.)	14.30	12.94	14.43	8.67	14.84	9.12
Female	36.38%	42.33%	37.32%	43.30%	40.32%	22.39%
Labour market outcomes						
Hourly wages	18.32	15.55	18.28	15.37	16.75	10.82
(std.dev.)	10.02	8.91	10.28	8.18	9.53	3.64
Number of						
person-year	24,494,734	3,517,484	18,299,923	1,288,395	13,124,303	116,794
persons	4,709,806	787,776	3,915,961	460,196	3,194,147	45,598
Firm level						
Firm size	60.96		115.80		214.59	
(std.dev.)	448.47		726.17		1,232.64	
Immigrant share	22.56%		32.46%		29.11%	
(std.dev.)	20.42%		24.34%		23.36%	
VA/worker	69.19		69.42		59.71	
(std.dev.)	38.75		41.27		34.64	
Number of firms	62,807		25,733		9,176	

Note: Hourly wages are deflated to provide 2015 prices . Firm size is measured as full-time equivalent employee units. Employers' characteristics are weighted by employment.

C.0.2 Age normalisation

Figure C.3: Wage profile against ages

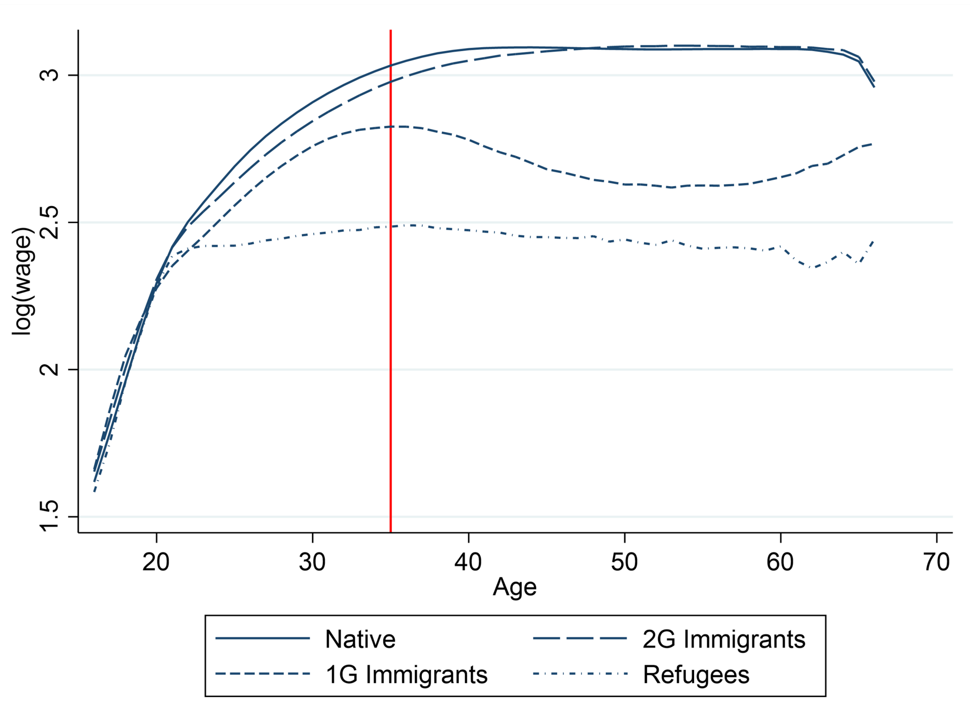


Table C.2: Mean age by nativity group

	Observation	Mean age
Natives	55,722,555	40.01
Second-generation immigrants	6,987,642	33.71
First-generation immigrants	2,481,491	31.25
Refugees	217,875	28.50
Simple mean age		33.37

Note: Simple mean age takes a simple mean of average ages of each nativity group (i.e., $\overline{age} = \sum_{g=1}^{4} \overline{age_g}$ where g is equal to 1 if native, 2 if second-generation immigrants, 3 if first-generation immigrants and takes a value of 4 if refugees).

C.0.3 Exogenous mobility

Table C.3: Wages of job changes for movers

Origin/	Number	Mean log wages				Changes
destination	changes	t_{-2}	t_{-1}	t_0	t_1	t_{-2} - t_1
quartile	(1)	(2)	(3)	(4)	(5)	(6)
Natives						
1 to 1	87,753	2.324	2.332	2.372	2.387	0.063
1 to 2	60,100	2.339	2.345	2.765	2.760	0.420
1 to 3	35,735	2.310	2.316	3.034	3.021	0.711
1 to 4	25,412	2.322	2.325	3.273	3.217	0.895
2 to 1	32,589	2.713	2.738	2.439	2.465	-0.248
2 to 2	107,918	2.782	2.794	2.801	2.796	0.014
2 to 3	57,335	2.801	2.806	3.024	3.009	0.208
2 to 4	38,292	2.770	2.778	3.294	3.250	0.480
3 to 1	11,690	3.007	3.029	2.432	2.466	-0.541
3 to 2	55,888	2.997	3.019	2.815	2.819	-0.178
3 to 3	98,764	3.029	3.040	3.042	3.035	0.006
3 to 4	68,513	3.058	3.062	3.281	3.255	0.197
4 to 1	10,147	3.195	3.308	2.417	2.449	-0.746
4 to 2	25,763	3.241	3.296	2.793	2.798	-0.444
4 to 3	52,475	3.242	3.272	3.063	3.062	-0.180
4 to 4	157,150	3.311	3.325	3.333	3.316	0.005
Second-generation immigrants						
1 to 1	5,405	2.187	2.202	2.255	2.278	0.091
1 to 2	6,017	2.224	2.236	2.625	2.624	0.400
1 to 3	4,614	2.215	2.224	2.965	2.955	0.739
1 to 4	3,108	2.215	2.225	3.261	3.213	0.998
2 to 1	2,093	2.579	2.602	2.317	2.343	-0.235
2 to 2	8,270	2.609	2.623	2.625	2.627	0.018
2 to 3	5,933	2.642	2.654	2.956	2.943	0.301
2 to 4	4,754	2.626	2.640	3.293	3.253	0.627
3 to 1	732	2.935	2.961	2.294	2.348	-0.587
3 to 2	4,378	2.927	2.948	2.677	2.688	-0.239
3 to 3	13,312	2.958	2.970	2.975	2.966	0.008
3 to 4	8,708	2.999	3.004	3.274	3.252	0.253
4 to 1	610	3.148	3.296	2.288	2.331	-0.817
4 to 2	2,246	3.223	3.292	2.653	2.661	-0.562
4 to 3	6,080	3.233	3.263	3.011	3.011	-0.222
4 to 4	19,168	3.307	3.321	3.329	3.314	0.007

Continued on next page

Origin/ destination quartile	Number changes (1)	Mean log wages t_{-2} (2)	t_{-1} (3)	t_0 (4)	t_1 (5)	Changes t_{-2} - t_1 (6)
First-generation immigrants						
1 to 1	718	2.230	2.245	2.264	2.274	0.044
1 to 2	575	2.290	2.303	2.497	2.492	0.202
1 to 3	682	2.297	2.311	2.809	2.801	0.504
1 to 4	159	2.252	2.270	3.209	3.100	0.849
2 to 1	343	2.443	2.488	2.292	2.307	-0.136
2 to 2	790	2.468	2.486	2.495	2.493	0.025
2 to 3	967	2.471	2.490	2.808	2.796	0.325
2 to 4	261	2.484	2.500	3.190	3.148	0.664
3 to 1	95	2.742	2.784	2.301	2.326	-0.416
3 to 2	320	2.723	2.755	2.510	2.521	-0.202
3 to 3	1,041	2.803	2.822	2.852	2.839	0.036
3 to 4	1,377	2.888	2.901	3.307	3.288	0.400
4 to 1	59	2.985	3.258	2.251	2.272	-0.713
4 to 2	84	3.052	3.191	2.518	2.534	-0.518
4 to 3	603	3.175	3.211	2.908	2.926	-0.249
4 to 4	3,038	3.263	3.282	3.361	3.349	0.086
Refugees						
1 to 1	21	2.010	2.011	2.092	2.141	0.132
1 to 2	16	1.978	2.037	2.384	2.337	0.359
1 to 3	10	2.039	2.082	2.548	2.527	0.488
1 to 4	31	2.008	2.060	2.864	2.811	0.803
2 to 1	8	2.352	2.373	2.108	2.055	-0.297
2 to 2	15	2.302	2.326	2.345	2.299	-0.003
2 to 3	15	2.297	2.350	2.533	2.528	0.231
2 to 4	27	2.306	2.350	2.773	2.730	0.424
3 to 3	9	2.489	2.525	2.521	2.497	0.008
3 to 4	15	2.530	2.552	2.888	2.777	0.247
4 to 1	5	2.660	2.756	2.132	2.140	-0.519
4 to 2	6	2.561	2.719	2.368	2.375	-0.187
4 to 3	7	2.649	2.678	2.558	2.527	-0.122
4 to 4	61	2.776	2.831	2.988	2.921	0.145

Note: Origin/destination quartiles are based on the mean wages of co-workers in the year before or the year after the job move. For confidentiality reasons, the mean wages of refugees who moved from the third quartile to the first and second quartiles are hidden.

C.0.4 Robustness check for AKM model

Table C.4: Summary of estimated AKM models with education

	Natives (1)	2nd-gen (2)	1st-gen (3)	Refugee (4)
Mean of log wages	2.883	2.729	2.640	2.329
Standard Deviation of log wages	0.517	0.545	0.435	0.334
Number of person-year obs.	48,316,915	5,849,906	1,776,790	132,739
Panel A: Summary of parameter estimates				
Number of person effects	7,219,653	955,339	479,279	38,927
Number of firm effects	479,769	172,031	51,923	13,222
Standard Deviation of person effect	0.389	0.356	0.324	0.286
Standard Deviation of firm effect	0.139	0.157	0.151	0.167
Standard Deviation of $X\beta$	0.253	0.280	0.149	0.406
Standard Deviation of residuals	0.133	0.154	0.134	0.114
Correlation of person and firm effects	0.220	0.120	0.133	-0.259
Correlation of person effect and $X\beta$	-0.099	0.048	0.025	-0.728
Correlation of firm effect and $X\beta$	0.162	0.238	0.133	0.132
Root MSE of model	0.145	0.172	0.159	0.145
Adjusted R^2 of model	0.922	0.901	0.866	0.813
Panel B: Share of variance of log-wages due to:				
Person effects + $X\beta$	0.731	0.721	0.684	0.696
Firm effects	0.072	0.083	0.121	0.250
Covariance of person and firm effects	0.089	0.045	0.069	-0.222
Covariance of firm and $X\beta$	0.042	0.070	0.032	0.160
Residuals	0.066	0.080	0.094	0.116

Note: The model is estimated separately for natives (Column (1)), second-generation immigrants (Column (2)), first-generation immigrants (Column (3)) and refugees (Column (4)) within the connected set of each group.

Table C.5: Summary of estimated AKM models with different flat ages

	Age 30				Age 40			
	Natives (1)	2nd-gen (2)	1st-gen (3)	Refugee (4)	Natives (5)	2nd-gen (6)	1st-gen (7)	Refugee (8)
Mean of log wages	2.882	2.729	2.650	2.342	2.882	2.729	2.650	2.342
Standard Deviation of log wages	0.515	0.541	0.439	0.331	0.515	0.541	0.439	0.331
Number of person-year obs.	54,555,580	6,710,024	2,281,716	183,628	54,555,580	6,710,024	2,281,716	183,628
Panel A: Summary of parameter estimates								
Number of person effects	7,407,587	1,003,001	576,689	48,279	7,407,587	1,003,001	576,689	48,279
Number of firm effects	513,499	185,902	59,757	16,470	513,499	185,902	59,757	16,470
Standard Deviation of person effect	0.643	0.573	0.359	0.189	0.306	0.293	0.320	0.303
Standard Deviation of firm effect	0.141	0.159	0.152	0.169	0.141	0.159	0.152	0.169
Standard Deviation of $X\beta$	0.416	0.337	0.166	0.260	0.330	0.388	0.179	0.415
Standard Deviation of residuals	0.136	0.158	0.139	0.120	0.136	0.158	0.139	0.120
Correlation of person and firm effects	0.262	0.222	0.182	-0.258	0.128	-0.009	0.173	-0.232
Correlation of person effect and $X\beta$	-0.743	-0.613	-0.254	-0.336	-0.066	-0.137	-0.072	-0.771
Correlation of firm effect and $X\beta$	-0.085	-0.019	0.088	0.146	0.285	0.318	0.138	0.143
Root MSE of model	0.147	0.174	0.163	0.149	0.147	0.174	0.163	0.149
Adjusted R^2 of model	0.918	0.897	0.862	0.797	0.918	0.897	0.862	0.797
Panel B: Share of variance of log-wages due to:								
Person effects + $X\beta$	0.713	0.699	0.655	0.641	0.713	0.699	0.655	0.641
Firm effects	0.075	0.086	0.119	0.260	0.075	0.086	0.119	0.260
Covariance of person and firm effects	0.179	0.138	0.103	-0.150	0.042	-0.003	0.087	-0.217
Covariance of firm and $X\beta$	-0.037	-0.007	0.023	0.117	0.100	0.133	0.039	0.183
Residuals	0.070	0.085	0.099	0.132	0.070	0.085	0.099	0.132

Note: The model is estimated separately for natives (Columns (1) and (5)), second-generation immigrants (Columns (2) and (6)), first-generation immigrants (Columns (3) and (7)) and refugees (Columns (4) and (8)) within the connected set of each group.

C.0.5 Normalisation of firm effect

Table C.6: Comparison goodness fit of models with different thresholds

	Adjusted R^2						Root MSE					
	Native & 2nd-gen		Native & 1st-gen		Native & Refugee		Native & 2nd-gen		Native & 1st-gen		Native & Refugee	
τ	Native	Immig	Native	Immig	Native	Immig	Native	Immig	Native	Immig	Native	Immig
3.00	0.1940	0.0910	0.2270	0.0785	0.1970	0.0178	0.1040	0.1370	0.1010	0.1530	0.0978	0.1820
3.10	0.1980	0.0934	0.2320	0.0810	0.2020	0.0192	0.1030	0.1370	0.1010	0.1530	0.0974	0.1820
3.20	0.2010	0.0959	0.2360	0.0836	0.2070	0.0207	0.1030	0.1370	0.1010	0.1530	0.0971	0.1820
3.30	0.2040	0.0979	0.2400	0.0859	0.2100	0.0218	0.1030	0.1360	0.1000	0.1530	0.0969	0.1820
3.40	0.2050	0.0992	0.2420	0.0882	0.2120	0.0226	0.1030	0.1360	0.1000	0.1520	0.0968	0.1820
3.50	0.2060	0.1000	0.2430	0.0904	0.2130	0.0231	0.1030	0.1360	0.1000	0.1520	0.0968	0.1820
3.60	0.2070	0.1020	0.2430	0.0929	0.2140	0.0239	0.1030	0.1360	0.1000	0.1520	0.0967	0.1810
3.70	0.2080	0.1040	0.2440	0.0959	0.2170	0.0262	0.1030	0.1360	0.1000	0.1520	0.0965	0.1810
3.75	0.2080	0.1050	0.2440	0.0971	0.2190	0.0279	0.1030	0.1360	0.1000	0.1520	0.0964	0.1810
3.80	0.2080	0.1060	0.2430	0.0980	0.2190	0.0297	0.1030	0.1360	0.1000	0.1520	0.0964	0.1810
3.85	0.2060	0.1060	0.2420	0.0986	0.2190	0.0315	0.1030	0.1360	0.1000	0.1520	0.0964	0.1810
3.90	0.2030	0.1060	0.2390	0.0987	0.2170	0.0335	0.1030	0.1360	0.1010	0.1520	0.0965	0.1810
3.95	0.2000	0.1050	0.2350	0.0982	0.2140	0.0356	0.1030	0.1360	0.1010	0.1520	0.0967	0.1800
4.00	0.1940	0.1030	0.2290	0.0969	0.2080	0.0375	0.1040	0.1360	0.1010	0.1520	0.0971	0.1800
4.05	0.1880	0.1000	0.2220	0.0948	0.2000	0.0389	0.1040	0.1360	0.1020	0.1520	0.0976	0.1800
4.10	0.1790	0.0960	0.2130	0.0913	0.1880	0.0390	0.1050	0.1370	0.1020	0.1520	0.0983	0.1800
4.20	0.1620	0.0882	0.1940	0.0801	0.1670	0.0360	0.1060	0.1370	0.1030	0.1530	0.0995	0.1800
4.30	0.1440	0.0796	0.1740	0.0682	0.1470	0.0325	0.1070	0.1380	0.1050	0.1540	0.1010	0.1810
4.40	0.1260	0.0705	0.1530	0.0574	0.1270	0.0289	0.1080	0.1380	0.1060	0.1550	0.1020	0.1810
4.50	0.1090	0.0617	0.1320	0.0475	0.1080	0.0259	0.1090	0.1390	0.1070	0.1560	0.1030	0.1810

Note: The above table presents the adjusted R^2 and root MSE of Equation 3.3 with different thresholds (τ).

Table C.7: Regressions of estimated firm effects on mean log value added per worker

	Native & 2nd-gen		Native & 1st-gen		Native & refugee	
	Native	Immig	Native	Immig	Native	Immig
	(1)	(2)	(3)	(4)	(5)	(6)
Slope	0.133***	0.118***	0.135***	0.118***	0.145***	0.0953***
	(4.93e-05)	(6.55e-05)	(5.41e-05)	(8.38e-05)	(7.59e-05)	(0.000155)
Constant	-0.0888***	-0.0726***	-0.0875***	-0.0627***	-0.0854***	-0.00133***
	(2.66e-05)	(3.53e-05)	(3.15e-05)	(4.89e-05)	(3.36e-05)	(6.61e-05)
Threshold (τ)	3.80	3.80	3.75	3.75	3.85	3.85

Note: The table displays coefficients from nonlinear regressions of non-normalised estimated firm effects for firms in dual-connected sets on Equation 3.3. Columns (1) and (2) are based on dual-connected sets of natives and second-generation immigrants, Columns (3) and (4) refer to the dual-connected set between natives and first-generation immigrants, and Columns (5) and (6) are the sets of natives and refugees.

C.0.6 Robustness check for decomposition model

Table C.8: Comparison of firms in dual-connected sets with low and high VA/L

	Native & 2nd-gen		Native & 1st$_{gen}$		Native & Refugee	
	Low VA/L	High VA/L	Low VA/L	High VA/L	Low VA/L	High VA/L
	(1)	(2)	(3)	(4)	(5)	(6)
Mean Firm Characteristics (Each firm weighted equally)						
Firm size	2.50	2.81	3.10	3.47	3.58	3.98
Immigrant %	15.6	11.9	16.2	10.2	4.5	2.6
Log wage	2.60	2.87	2.59	2.90	2.50	2.82
Log sales/L	4.53	5.42	4.40	5.46	4.43	5.44
Industry distribution (% of total firm number in low/high VA/L)						
Hospitality industry	17.1	3.7	23.3	4.9	25.8	6.2
Retail	16.2	7.7	10.3	5.0	12.6	6.2
Business services	10.8	15.2	11.2	17.0	5.3	9.1
Lending companies	5.6	1.0	11.0	1.9	12.8	1.7

Note: Sample includes firms in dual-connected sets with financial data only. High VA/L is defined as above the thresholds with Equation 3.3 (i.e. 3.80 and 3.75 for dual-connected sets between natives and second-generation immigrants and between natives and first-generation immigrants. 3.85 for the set of natives and refugees). Firm size is measured by the log of employment.

Figure C.4: Firm effects versus sales per worker

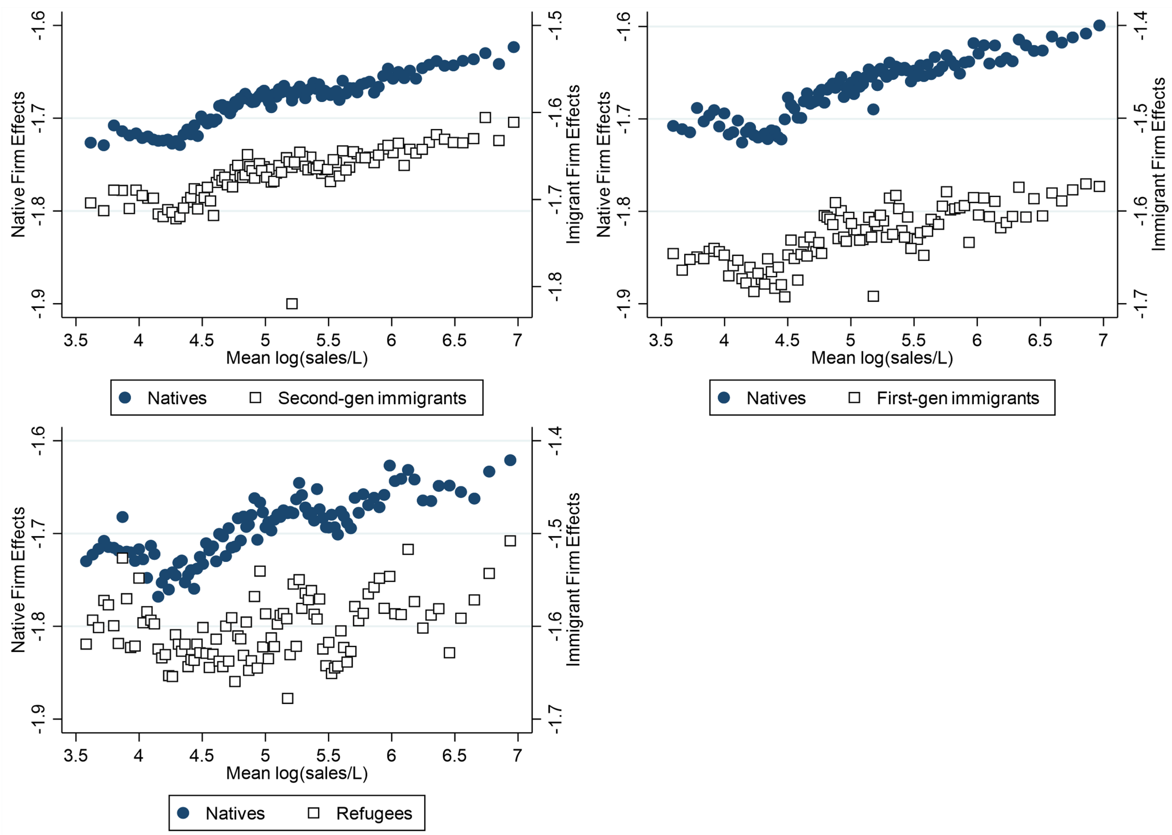


Table C.9: Regressions of estimated firm effects on mean log(sales/L)

	Native & 2nd-gen		Native & 1st-gen		Native & refugee	
	Native (1)	Immig (2)	Native (3)	Immig (4)	Native (5)	Immig (6)
Slope	0.0749***	0.0619***	0.0845***	0.0755***	0.100***	0.0756***
	(5.69e-05)	(7.30e-05)	(6.50e-05)	(9.68e-05)	(8.91e-05)	(0.000174)
Constant	-0.0470***	-0.0337***	-0.0412***	-0.0221***	-0.0496***	0.0276***
	(2.46e-05)	(3.16e-05)	(2.87e-05)	(4.29e-05)	(3.21e-05)	(6.05e-05)
Threshold (τ)	5.40	5.40	5.40	5.40	5.50	5.50

Note: The table displays coefficients from nonlinear regressions of non-normalised estimated firm effects for firms in dual-connected sets on Equation 3.3. Columns (1) and (2) are based on dual-connected sets of natives and second-generation immigrants, Columns (3) and (4) refer to the dual-connected set between natives and first-generation immigrants, and Columns (5) and (6) are the sets of natives and refugees.

Table C.10: Contribution of firm-specific pay premiums - Sales/L normalisation

	Pay gap	Mean firm premium		Total contribution	Sorting	Pay-setting
		Native	Immig	of firm effect	effect	effect
	(1)	(2)	(3)	(4)	(5)	(6)
Native - 2nd-gen	0.162	0.021	-0.005	0.026	0.003	0.023
				(0.16)	(0.02)	(0.14)
Native - 1st-gen	0.163	0.024	-0.002	0.026	-0.004	0.030
				(0.16)	(-0.03)	(0.19)
Native - Refugee	0.314	0.019	-0.119	0.139	0.034	0.105
				(0.44)	(0.11)	(0.33)

Note: Column (1) shows the mean log wage gap between natives and immigrants, in the dual-connected set. Columns (2) and (3) show the mean normalised pay premiums by groups. Column (4) is the difference between columns (2) and (3), referring to the total contribution of firm-specific sorting and pay-setting policies to the native-immigrant gap. Columns (5) and (6) decompose the total in column (4) into firm sorting effects (5) and pay-setting effects (6). Entries in parentheses are the shares of the overall gap explained by the component in the column.