FACULTY OF SOCIAL SCIENCES
Department of Economics
University of Copenhagen

Essays in Migration Economics

A Host Country Perspective on Receiving and Integrating Immigrants

Ph.D. Dissertation

Linea Hasager

Advisors: Jakob Roland Munch & Mette Foged

Submitted: August 31, 2020

Ph.D. Dissertation

Essays in Migration Economics

A Host Country Perspective on Receiving and Integrating Immigrants

Linea Hasager
Department of Economics
University of Copenhagen

Submitted: August 31, 2020
Advisors: Jakob Roland Munch & Mette Foged

Contents

Acknowledgements

The past three years have been a fun and challenging experience. Writing this Ph.D. dissertation would not have been possible without the generous financial support from the Innovation Fund Denmark.

Above all, I wish to thank my advisor and co-author Mette Foged, whom provided me with this opportunity. She has been crucial for this project coming to light. My advisor Jakob Roland Munch provided continuous support and guidance, for which I am grateful.

I am indebted to my co-authors Vasco Yasenov, Giovanni Peri, Jacob Nielsen Arendt and Iben Bolvig for their dedicated work. I would like to extend a special thanks to my co-author and good friend Mia Jørgensen for the endless conversations on Skype discussing the tiniest details, while overcoming the 9-hour time difference between California and Copenhagen for several months.

Giovanni Peri gave me the opportunity to spend six months at the University of California, Davis. I wish to thank him for this, and I want to thank faculty and Ph.D. students at UC Davis for their valuable comments.

My fellow Ph.D. students at the University of Copenhagen made this journey much more pleasant. They deserve thanks for discussions in the office, lunch breaks on the balcony, coffee runs, Friday beers and the Thursday running club.

Finally, I thank my boyfriend Mikkel for his support and for always being eager to share his knowledge on data issues, econometrics and life as a Ph.D. student. Thank you for making working from home during the pandemic in the spring of 2020 a “not-so-terrible” experience.

Linea Hasager
August 31, 2020

Introduction

This Ph.D. dissertation consists of four self-contained chapters. Although the chapters study different research questions and the applied methods differ, they all share a common focus on how host countries absorb immigrants into society.

In the first chapter we study the health consequences of living in a low-income neighborhood. To do so, we exploit the nature of a Spatial Dispersal Policy which resettled refugees quasi-randomly across Danish neighborhoods. This setting allows us to separate the causal effect of living in the poorest neighborhoods from any selection into neighborhoods. We show that living in the poorest third of neighborhoods significantly increases the risk of developing a range of lifestyle related diseases such as diabetes type II, obesity and hypertension, and these impacts are more pronounced for women and younger individuals. Our findings suggest that the development of these lifestyle related diseases is caused by interaction with immediate neighbors along with the characteristics of the small geographical area, whereas differences in individual income, access to health care and the presence of ethnic networks appear less important.

The second chapter is concerned with how language training and economic work incentives affect refugees' long-term economic and social integration. We compare long-run outcomes for refugees admitted to Denmark around a reform in 1999. This reform mandated a significant increase in language learning instruction for those who arrived just after January 1, 1999. A subgroup of these refugees also experienced a one-year temporary reduction in welfare benefits. The expanded language training resulted in a significant increase in employment and earnings over the following 18 years, while the temporary benefit reduction had no impact on labor market outcomes. However, the reduced welfare benefits caused crime rates to increase in the first year due to a rise in shoplifting committed in supermarkets. Moreover, the improvements in language proficiency and labor market outcomes had positive spillovers on the next generation as male children of the treated refugees were more likely to complete secondary schooling and were less likely to commit crime.

In the third chapter I study how integration of female refugees is affected by their type of residence permit. My findings show that women are more likely to be admitted through family-reunification procedures than asylum recognition compared to men. However, if women are recognized as refugees, their employment and earnings gradually increase over time, while divorce rates increase and the risk of being subject to intimate partner violence decreases immediately after asylum recognition. The findings from my empirical analysis are consistent with the predictions from a theoretical model of household bargaining, where a decrease in the woman's risk of being returned to her origin country increases her share of household consumption and decreases the level of domestic violence.

The fourth chapter documents the role of labor market institutions in affecting the wage impacts of immigration. By gathering a database of previously reported semi-elasticities from academic studies, we show that immigration in general has small positive impacts on natives' average wages, but relative wages decline for the workers with skills most similar to the immigrants. Our findings suggest that the decline in relative wages is mitigated by higher labor market rigidity, which implies that institutions can reduce the distributional consequences of immigration. However, this comes at the expense of reducing the potential overall gain from immigration.

For convenience the four abstracts are listed below.

Part 1: Sick of Your Poor Neighborhood? Quasi-Experimental Evidence on Neighborhood Effects in Health

with Mia Herløv Jørgensen

Does living in a low-income neighborhood have negative health consequences? We document neighborhood effects in health by exploiting a Spatial Dispersal Policy that resettled refugees quasi-randomly across neighborhoods from 1986 to 1998, which allows us to separate causal impacts from selection into neighborhoods. We show that the risk of developing a lifestyle related disease before 2018 increases by 5.4 percent relative to the sample mean for individuals who were allocated to the poorest third of neighborhoods compared to allocation to the richest third of neighborhoods. In particular, among women and younger individuals the impact of neighborhood income on health is larger. Differences across neighborhoods in access to health care, presence of ethnic networks, and individual labor market outcomes – and thus also individual income growth differences – cannot explain our findings. Instead, our results suggest that interaction with immediate neighbors and the characteristics of the very local environment are important for understanding neighborhood effects in health, especially when considering diabetes and obesity.

Part 2: Integrating Refugees: The Role of Language Training and Work Incentives

with Jacob Nielsen Arendt, Iben Bolvig, Mette Foged & Giovanni Peri

Social and economic integration of refugees are key to their personal success and to producing positive effects in the host country. We evaluate the effects of a reform that substantially expanded language training for immigrants who obtained refugee status in Denmark on or after January 1, 1999. The same reform also temporarily decreased welfare benefits for a subgroup of them. Using a regression discontinuity design around the cutoff date we find positive and significant employment and earnings effects on the treated group, relative to the untreated group. Employment increased by 23 percent (4 percentage points) and yearly earnings increased by 34 percent (USD 2,500) when measured eighteen years after the start of the language program. We do not find any labor market effect of the reduction of welfare benefits. We find, however, evidence of temporarily higher property crimes for the group subject to lower benefits. The labor market effects are much stronger for individuals with Arabic/Dari mother language, consistently with a more crucial role of language training for speakers of languages that are very different from Danish. Finally, male children of treated refugees were more likely to complete lower secondary school and less likely to commit crime.

Part 3: Does Granting Refugee Status to Family-Reunified Women Improve Their Integration?

In many refugee-receiving countries men are the principal asylum applicant, while women are admitted through family-reunification procedures. I document that granting asylum to family-reunified women has significant impacts on economic integration and decreases their risk of being victims of intimate partner violence. Using an event study approach, I find that annual employment and earnings increase by 1.6 percentage points and USD 500, respectively, immediately after asylum recognition relative to family-reunification status. These are large effects compared to the low baseline of virtually zero employment and earnings in the preceding years. At the same time the divorce rate increases by 3.8 percentage points and domestic violence decreases by 0.9 percentage points. The decrease in violence is observed regardless of whether the woman remains married or not, which suggests that the new, more favorable, residence permit improves her bargaining power within the marriage. This is consistent with the predictions from a Nash bargaining model where the risk of being returned to the home country affects the woman's threat point, and thus the allocation of resources within the marriage.

Part 4: Labor Market Rigidities and the Wage Impact of Immigration

with Mette Foged & Vasil Yasenov

We study the role of labor market institutions in affecting the wage impact of immigration using a cross-country meta-analysis approach. We gather information on 1,030 previously reported semi-elasticities from 54 academic studies covering 18 developed countries. We supplement this dataset with country-level institutional strength and coverage data from

the OECD. Our results suggest that higher labor market rigidity mitigates the effect on relative wages of native workers with skills most similar to immigrants but exacerbates the impacts on average earnings in the economy. In other words, institutions shield native workers from distributional consequences but diminish potential benefits induced by foreign labor.

Introduktion

Denne ph.d.-afhandling består af fire selvstændige kapitler. På trods af at kapitlerne omhandler forskellige forskningsspørgsmål og anvender forskellige metoder, har de alle det tilfældes, at de fokuserer på, hvordan immigranter absorberes i modtagerlandet.

I det første kapitel studerer vi de helbredsmæssige konsekvenser af at bo i et lavindkomst-nabolag. Her udnytter vi en fordelingspolitik, som genbosatte flygtninge kvasi-tilfældigt i forskellige danske nabolag. Denne fordelingspolitik gør det muligt at adskille den kausale effekt af at bo i de fattigste nabolag fra selektion til nabolag. Vores resultater viser, at allokering til den fattigste tredjedel af nabolag reducerer risikoen for at udvikle en række livsstilsygdomme såsom diabetes type II, overvægt og hypertension, og disse effekter er størst for kvinder og yngre personer. Desuden antyder vores resultater, at disse livsstilsygdomme skyldes interaktion med umiddelbare naboer samt det lokale nabolags karakteristika, mens forskelle i individuel indkomst, adgang til sundhedstilbud samt tilstedeværelsen af etniske netværk synes mindre vigtige i denne sammenhæng.

Andet kapitel undersøger hvordan sprogtræning og økonomiske incitamenter påvirker flygtninges økonomiske og sociale integration på lang sigt. I dette kapitel sammenlignes langsigtede resultater for immigranter i Danmark, som blev anerkendt som flygtninge omkring en reform af det danske integrationsprogram i 1999. Denne reform resulterede i en markant forøgelse af sprogundervisning for flygtninge, som ankom efter 1. januar 1999. En subpopulation af disse flygtninge oplevede desuden en midlertidig ydelsesreduktion i løbet af det første år. Den udvidede sprogundervisning resulterede i en signifikant stigning i beskæftigelsen og lønninger i de efterfølgende 18 år, mens den midlertidige ydelsesreduktion ikke påvirkede hverken løn eller beskæftigelse. Dog resulterede ydelsesreduktionen i mere kriminalitet i det første år grundet tyveri i supermarkeder. Derudover havde de bedre sprogforudsætninger samt øget løn og beskæftigelse en positiv spillover-effekt på den næste generation. Sønner af voksne flygtninge, som modtog øget sprogundervisning, havde større sandsynlighed for at færdiggøre en ungdomsuddannelse, og de var mindre tilbøjelige til at begå kriminalitet.

I det tredje kapitel studerer jeg, hvordan integration af kvindelige flygtninge påvirkes af deres type af opholdstilladelse. Mine resultater viser, at kvinder har markant højere sandsynlighed for at have familiesammenføringsstatus end flygtningestatus sammenlignet med mænd. Hvis familiesammenførte kvinder i stedet anerkendes som flygtninge, betyder det, at deres løn og beskæftigelse stiger gradvist over tid, mens skilsmisseraten stiger og risikoen for at blive udsat for vold i hjemmet reduceres umiddelbart efter, at kvinden anerkendes som flygtning. Resultaterne fra den empiriske analyse er konsistente med forudsigelserne fra en teoretisk model, der beskriver en husholdnings forhandling vedrørende forbrug og vold i hjemmet. Modellen tilsiger, at en lavere risiko for, at kvinden sendes tilbage til sit hjemland, vil betyde, at hendes andel af husholdningens forbrug øges, mens vold i hjemmet reduceres.

Det fjerde kapitel dokumenterer sammenhængen mellem arbejdsmarkedsinstitutioner og løneffekter af indvandring. Baseret på en database af tidligere rapporterede semi-elasticiteter fra akademiske artikler viser vores resultater, at indvandring generelt har en lille positiv effekt på gennemsnitlige lønninger i modtagerlandet, mens relative lønninger falder for de arbejdere i den oprindelige arbejdsstyrke, som ligner immigranterne mest. Endvidere viser vores resultater, at faldet i de relative lønninger modvirkes af et mere rigidt arbejdsmarked, hvilket betyder, at stærke arbejdsmarkedsinstitutioner kan reducere de fordelingsmæssige konsekvenser af immigration. Dog sker dette på bekostning af en reduktion i de overordnede potentielle gevinster ved indvandring.

De fire abstracts er indsat herunder.

Part 1: Sick of Your Poor Neighborhood? Quasi-Experimental Evidence on Neighborhood Effects in Health

med Mia Herløv Jørgensen

Er der negative sundhedskonsekvenser forbundet med at bo i et lavindkomst-nabolag? Vi dokumenterer nabolagseffekter i sundhed ved at udnytte en fordelingspolitik, hvor flygtninge blev genbosat kvasi-tilfældigt i forskellige nabolag fra 1986 til 1998, hvilket gør os i stand til at separere kausale effekter fra selektion til nabolag. Vores resultater viser, at risikoen for at udvikle en livsstilssygdom før 2018 øges med 5,4 procent af populations gennemsnit ved allokering til den fattigste tredjedel af nabolag i forhold til allokering til den rigeste tredjedel af nabolag. Disse effekter er størst blandt kvinder samt yngre personer. Vores resultater kan ikke forklares af forskelle i adgang til sundhedstilbud på tværs af nabolag, tilstedeværelsen af etniske netværk eller individuelle løn- og beskæftigelsesforhold – herunder individuel indkomstudvikling. I stedet antyder vores resultater, at interaktion med de nærmeste naboer samt det umiddelbare lokalområdes karakteristika har en vigtig betydning for nabolagseffekter i sundhed – især når det drejer sig om diabetes og overvægt.

Part 2: Integrating Refugees: The Role of Language Training and Work Incentives

med Jacob Nielsen Arendt, Iben Bolvig, Mette Foged & Giovanni Peri

Social og økonomisk integration af flygtninge er en vigtig forudsætning for deres personlige succes samt for at generere positive effekter for modtagerlandet. Vi evaluerer effekterne af en reform, som forøgede sprogtræning markant for immigranter, der blev anerkendt som flygtninge i Danmark den 1. januar 1999 eller derefter. Denne reform introducerede desuden en midlertidig ydelsesreduktion for en subpopulation af de nyankomne flygtninge. Vi finder positive signifikante effekter på beskæftigelse og løn for flygtninge, der modtog mere sprogundervisning, relativt til kontrolgruppen ved brug af et "Regression Discontinuity Design" omkring den

1. januar 1999. Beskæftigelsen steg med 23 percent (4 procentpoint) og årlig indkomst blev forøget med 34 procent (USD 2.500) målt atten år efter starten af integrationsprogrammet. Vi finder ingen effekter på løn og beskæftigelse af den midlertidige ydelsesreduktion. Derimod finder vi evidens for et midlertidigt højere antal ejendomsforbrydelser for den delmængde af populationen, der modtog den lavere ydelse. Arbejdsmarkedseffekterne er stærkere for personer med arabisk eller dari som modersmål, hvilket er konsistent med, at sprogtræning er vigtigere for personer, der taler sprog, som er meget forskellige fra dansk. Derudover dokumenterer vi, at sønner af voksne flygtninge, som modtog mere danskundervisning, var mere tilbøjelige til at færdiggøre en ungdomsuddannelse samt mindre tilbøjelige til at begå kriminalitet.

Part 3: Does Granting Refugee Status to Family-Reunified Women Improve Their Integration?

I mange lande, der modtager flygtninge, er hovedansøgeren typisk en mand, mens kvinder oftest ankommer ved hjælp af familiesammenføring. Jeg dokumenterer, at der er signifikante positive effekter på økonomisk integration samt en reduktion i risikoen for, at kvinderne bliver udsat for vold i hjemmet, hvis de anerkendes som flygtninge i stedet for at have familiesammenføringsstatus. Ved brug af en "Event Study"-tilgang finder jeg, at årlig beskæftigelse og løn stiger med henholdsvis 1,6 procentpoint og USD 500 umiddelbart efter, at kvinderne anerkendes som flygtninge. Dette er forholdsvist store effekter med henblik på det lave niveau i årene op til, hvor de stort set ingen beskæftigelse eller indkomst har. Samtidig stiger skilsmisseraten med 3,8 procentpoint, mens risikoen for vold i hjemmet reduceres med 0,9 procentpoint. Denne reduktion i risikoen for at blive udsat for vold findes både for kvinder, der skilles fra deres mænd samt for kvinder, der fortsætter ægteskabet. Dette antyder, at kvindens nye favorable opholdstilladelse forbedrer hendes forhandlingskraft i ægteskabet. Dette er konsistent med prædiktionerne fra en Nash Bargaining Model, hvor sandsynligheden for at blive sendt tilbage til hjemlandet påvirker kvindens "threat point", og derved fordelingen af ressourcer inden for ægteskabet.

Part 4: Labor Market Rigidities and the Wage Impact of Immigration
med Mette Foged & Vasil Yasenov
Vi studerer arbejdsmarkedsinstitutioners sammenhæng med indvandringens løneffekter ved hjælp af en metaanalyse på tværs af lande. Vi har indsamlet information vedrørende 1.030 tidligere estimerede semi-elasticiteter fra 54 akademiske studier, som dækker 18 udviklede lande. Dette datasæt suppleres med data fra OECD vedrørende institutioners styrke og dækning på landeniveau. Vores resultater viser, at den negative effekt på relative lønninger for de arbejdere, som ligner immigranterne mest, modvirkes af et mere rigidt arbejdsmarked, mens den positive effekt på gennemsnitlige lønninger dæmpes. Med andre ord kan arbejdsmarkedsinstitutioner beskytte den oprindelige arbejdsstyrke mod fordelingsmæssige konsekvenser af indvandring, men samtidig reducerer et rigidt arbejdsmarked de potentielle gevinster ved udenlandsk arbejdskraft.

Chapter 1

Sick of Your Poor Neighborhood? Quasi-Experimental Evidence on Neighborhood Effects in Health

Sick of your poor neighborhood? *

Quasi-experimental evidence on neighborhood effects in health

Linea Hasager† Mia Jørgensen‡

August 11, 2020

Abstract

Does living in a low-income neighborhood have negative health consequences? We document neighborhood effects in health by exploiting a Spatial Dispersal Policy that resettled refugees quasi-randomly across neighborhoods from 1986 to 1998, which allows us to separate causal impacts from selection into neighborhoods. We show that the risk of developing a lifestyle related disease before 2018 increases by 5.4 percent relative to the sample mean for individuals who were allocated to the poorest third of neighborhoods compared to allocation to the richest third of neighborhoods. In particular, among women and younger individuals the impact of neighborhood income on health is larger. Differences across neighborhoods in access to health care, presence of ethnic networks, and individual labor market outcomes – and thus also individual income growth differences – cannot explain our findings. Instead, our results suggest that interaction with immediate neighbors and the characteristics of the very local environment are important for understanding neighborhood effects in health, especially when considering diabetes and obesity.

JEL Classification: J15, I12, I14, I31
Keywords: Health inequality, Refugee Dispersal Policy, lifestyle related diseases, neighborhood effects

*We thank our advisors Asger Lau Andersen, Mette Foged, Niels Johannesen and Jakob Roland Munch for helpful comments and discussions. The project benefited from comments from Pernille Plato Jørgensen, Ida Lykke Kristiansen, participants at the Applied Micro Brown Bag Seminar at UC Davis, the Brown Bag Seminar at VIVE, the Ph.D. seminar and the CEBI Lunch at the University of Copenhagen. Linea acknowledges support from the Economic Assimilation Research Network (EARN), at the University of Copenhagen, financed by the Innovation Fund Denmark (grant #6149-00024B). Mia acknowledges support from the Center of Economic Behavior and Inequality (CEBI), at the University of Copenhagen, financed by grant #DNRF134 from the Danish National Research Foundation. All errors are our own.

†Department of Economics, University of Copenhagen, Øster Farimagsgade 5 Building 26, DK 1353 Copenhagen, tlh@econ.ku.dk.

‡Department of Economics, University of Copenhagen, Øster Farimagsgade 5 Building 26, DK 1353 Copenhagen, mia.joergensen@econ.ku.dk

1 Introduction

It is well established that affluent individuals on average live longer and have better health than those at the bottom of the income distribution.[1] It is therefore unsurprising that people living in higher income areas also have better health, on average.[2] However, richer areas might also provide substantially different types of amenities and display different social norms related to lifestyle choices, which in itself could improve residents' health. In this paper, we document significant causal impacts on a wide range of lifestyle related diseases that have not been studied previously, and, to the best of our knowledge, we are the first to explore the potential mechanisms behind neighborhood effects in health.[3]

The scarcity of knowledge on causal neighborhood effects in health may stem from the fact that establishing a causal relationship between residential location and health is notoriously difficult. One could ask if low-income residents are less healthy because they live in low-income areas? Or do they live in a low-income area due to their poor health which is also affecting their income? One could even imagine that individual income determines both neighborhood choice and health, and thus explains the observed neighborhood income gradient in health.

In other words, observing that residents in poorer areas have worse health does not necessarily imply that neighbors or the characteristics of the local area actually affect residents' health. Thus, to identify neighborhood effects in health the researcher needs variation in neighborhood income levels across individuals which is uncorrelated with their personal characteristics. Moreover, in order to separate long-term neighborhood income effects from individual income effects, individual income growth must be independent of the neighborhood's income level.

Fortunately, the nature of the natural experiment, we study, enables us to overcome the challenge of self-selection into neighborhoods in a setup where we can separate the impact of neighborhood income from individual income. We exploit a Danish Spatial Dispersal Policy in act from 1986 to 1998 that quasi-randomly assigned refugee families to different neighborhoods upon arrival to Denmark.[4] We define a neighborhood as a parish, which historically has delineated a small community and in recent years has been home to around 3,000 inhabitants. We divide all neighborhoods into three equally sized groups based on the median household income in the neighborhood one year prior to the refugees'

[1] See for example Chetty et al. (2016) who document the association between individual income and life expectancy in the United States.

[2] See for example Panels a-f in Figure A.1 in Appendix, which shows a negative correlation between median local area (parish) income and the share of inhabitants diagnosed with a number of different lifestyle related diseases.

[3] See Ludwig et al. (2011) and Ludwig et al. (2012) for existing evidence on obesity, elevated blood sugar levels and subjective well-being.

[4] A number of existing papers study this natural experiment. See Damm (2009), Damm and Dustmann (2014), Foged and Peri (2015) and Dustmann, Vasiljeva, and Damm (2018) among others.

arrival, and we document that the income group of the neighborhood, in which the refugee was placed by the Spatial Dispersal Policy, causally impacts the newcomers' health.

Our analysis is comprised of two different parts. First, we show that living in a low-income area increases the risk of developing several lifestyle related diseases. Being assigned to the poorest third of neighborhoods increases the risk of suffering from a lifestyle related disease by 5.4 percent relative to middle- or top-income neighborhoods, respectively. We find no significant impact on average mental health outcomes. Using an instrumental variables strategy, we find that for each year spent in the poorest third of neighborhoods, the risk of developing a lifestyle related disease increases by 0.3 percentage points.[5] This is primarily due to an increased risk of developing hypertensive diseases along with endocrine and nutritional diseases such as diabetes and obesity. Moreover, we show that the negative health effects of being assigned to the poorest third of neighborhoods are larger for young individuals below age 26 and females.[6]

In the second part of our analysis we take a step towards understanding the documented neighborhood income gradient in health. A neighborhood may influence its residents' physical and mental health in multiple ways. For example through transmission of behavior (e.g. health habits), its local amenities (e.g. recreational areas or grocery store options), labor market opportunities, or through the local institutions such as health care access. We study the risk of developing lifestyle related diseases since this risk strongly depends on individual behaviors that might be influenced by the factors mentioned above.[7] Since some of these factors may also affect mental health, we include mental health diagnoses in our analysis.[8] Moreover, lifestyle related diseases are an important health outcome, since these types of diseases are responsible for more than 70 percent of deaths worldwide each year (World Health Organization (2018)).

Interestingly, the estimated income gradient in health is not a result of more advantageous labor market outcomes for individuals placed in higher income neighborhoods. Our results consistently show that there are no significant differences in any labor market outcome across neighborhood income levels. This finding is in line with previous work studying neighborhood effects, which documents that there is no association between a local area's quality and labor market outcomes for residents (see Damm (2014), Sanbonmatsu et al. (2011), Kling, Liebman, and Katz (2007) or Oreopoulos (2003) among others).[9] We

[5]We instrument the number of years spent in the poorest third of neighborhoods by the initial placement neighborhood.

[6]This is consistent with previous evidence showing that the young are more susceptible to disadvantageous environments, see for example Kling, Liebman, and Katz (2007).

[7]See Patienthåndbogen (2017).

[8]We refer to Sanbonmatsu et al. (2011) for a complete overview of potential mechanisms through which neighborhoods may influence mental and physical health.

[9]Damm (2014) documents that refugees located in socially deprived areas do not experience worse labor market outcomes

can therefore rule out any income effects of neighborhood placement, and this allows us to attribute the estimated health effects to neighborhood income rather than to individual income.[10]

Next, we show that controlling for a number of neighborhood characteristics and resident composition does not affect the income gradient. The universal health care system in Denmark ensures that in general any differences in access to and quality of health care across geographical areas are small. Including additional controls for health care access and general health conditions in the municipality also leaves the income gradient in health unaffected. Furthermore, controlling for institutional differences between municipalities, differences between rural and urban parishes as well as the presence of ethnic networks does not affect the income gradient. Thus, these mechanisms do not appear to be important determinants of neighborhood health.

There are some mechanisms that we cannot measure and test directly. These are factors such as health behaviors of peers and local amenities. However, we take a step in that direction by documenting the importance of the very local environment. We do this by varying the level of a neighborhood using both a more aggregated level (municipalities) and a more disaggregated level (households living in same apartment complex), which changes how well we capture potential peer groups and the characterization of the immediate neighborhood. When we compare the resulting income gradients from these estimations, we find that the closer we get to immediate neighbors (and the very local geographical area in which the refugees live), the larger the estimated coefficients become. This suggests that transmission of behaviors from neighbors and local amenities are part of the mechanisms through which neighborhoods affect residents' health.

We base our analysis on administrative registers covering 31 years, which allows us to observe annual residential locations, income, hospital diagnoses and other individual characteristics. In spite of the high quality of our data, it is likely that our estimates are a lower bound of the true effect size due to varying detection rates across areas. Correlational evidence shows that a larger share of residents in richer neighborhoods visit their GP or dentist in a given year, see Panels g-h in Figure A.1 in Appendix.[11] This may result in lower detection rates in poorer neighborhoods which will lead to a downward bias in our estimates.

than those placed elsewhere. Similarly, the randomized controlled trial *"Moving To Opportunity"* literature does not suggest any long term effects on labor market attachment, economic self-sufficiency or income levels, see Sanbonmatsu et al. (2011) and Kling, Liebman, and Katz (2007).

[10] We show that in richer neighborhoods, more refugees obtain a vocational education, but there is no significant difference in the share obtaining a health-related education across neighborhoods. In addition, there are no differences in the task complexity of occupations, conditional on employment. Thus, any effects of education on health are most likely indirect effects such as more general knowledge, higher self-esteem or better working conditions.

[11] See Bago d'Uva and Jones (2009) for evidence on the association between health care utilization and income.

An important contributor to the knowledge on neighborhood effects has been the randomized controlled trial *Moving to Opportunity* experiment, which was carried out from 1994 to 1998 in five big American cities, see for example Katz, Kling, and Liebman (2001), Kling, Liebman, and Katz (2007) or Chetty, Hendren, and Katz (2016). However, because of data limitations the *Moving to Opportunity* experiment only provides limited evidence on neighborhood effects in health. The experiment shows that moving to a low-poverty neighborhood significantly improves subjective well-being (Ludwig et al. (2012)), decreases the risk of extreme obesity and elevated blood sugar levels (Ludwig et al. (2011)), and improves adult mental health (Kling, Liebman, and Katz (2007)).[12]

The literature also includes non-experimental evidence on neighborhood effects in health, for example in mental health among social housing clients, in life expectancy among elderly, and in diabetes among refugees (see Boje-Kovacs, Greve, and Weatherall (2018)[13], Finkelstein, Gentzkow, and Williams (2019)[14], and White et al. (2016)[15], respectively).

We contribute to the literature on neighborhood effects in health in two ways. Since causal evidence on neighborhood effects in health is limited, the first part of our contribution is to document the existence of strong and significant causal long-term neighborhood effects in a wide range of lifestyle related diseases. Previous work has studied diabetes, obesity, subjective well-being and mental health. We consider a number of medical conditions that have not been examined previously, and we are the first to quantify the impact of one additional year spent in a low-income neighborhood. Furthermore, since previous studies do not provide knowledge on the mechanisms, the second part of our contribution is to push forward the understanding of these neighborhood effects by ruling out a number of likely mechanisms and pointing to the importance of the nature of very local environments.

Because of this finding, our paper also relates to the literature on spillovers in health within smaller networks. This includes for example Eisenberg et al. (2013) who find no or small contagious effects of mental health between college roommates, Christakis and Fowler (2007) who document an increased risk of obesity within social networks if a person in that network becomes obese, and Fadlon and Nielsen (2019) who find spillovers in health behaviors among family members and coworkers.

In the remainder of the paper we first describe the Spatial Dispersal Policy that dispersed individu-

[12]Ludwig et al. (2012) also document non-significant improvements in two indices of mental and physical health. In Ludwig et al. (2011) elevated blood sugar level is included as an indication of untreated diabetes.

[13]Boje-Kovacs, Greve, and Weatherall (2018) study the impact on mental health of living in a socially deprived neighborhood for vulnerable residents in the capital of Denmark. They find an impact on mental health based on purchases of psychotropics (anti-depressants, anti-psychotics etc.).

[14]Finkelstein, Gentzkow, and Williams (2019) show that moving to a neighborhood with higher life expectancy increases the newcomer's life expectancy among Medicare recipients in the US by comparing movers from the same origin.

[15]White et al. (2016) show that neighborhood deprivation increases the risk of developing diabetes using a Swedish refugee dispersal policy similar to the one we use for identification.

als quasi-randomly to Danish neighborhoods, which lays the foundation for our identification strategy, Section 2. We carefully spell out the identifying assumptions, discuss potential threats to identification and provide balancing tests supporting the identifying assumptions in this section. Then we present our empirical model in Section 3, describing a reduced form approach and an instrumental variables strategy. In Section 4 we describe the data sources, sample selection and the definition of our main variables of interest. Following that, Section 5 provides an overview of our results which shows an increased risk of developing lifestyle related diseases as a consequence of living in a low-income neighborhood. In Section 6 we test a number of potential mechanisms and show the importance of the very local environment. Finally, in Section 7 we discuss the external validity of our findings, and Section 8 concludes the paper.

2 Institutional background and identification

2.1 The Danish Spatial Dispersal Policy, 1986 to 1998

From 1986 to 1998 the Danish Refugee Council (DRC) was in charge of Danish integration efforts targeted at newly arrived refugees. Among other things, this meant that the DRC was responsible for finding permanent housing for refugees. Prior to 1986 refugees had mainly found housing in the largest cities, but in 1986 the DRC adopted a Spatial Dispersal Policy (SDP) designed to spread refugees evenly across Denmark. In this section we highlight the features of the policy that created exogenous variation in the allocation of refugees across municipalities, parishes and stairways of apartment buildings.

Once the Danish government had granted asylum to an asylum seeker, the newly recognized refugee filled out a questionnaire with some basic information on age, ethnicity and family size.[16] We will refer to this information as 'questionnaire observables'. This questionnaire contained all information about the refugee that was available to the DRC at the time of allocation. The DRC used the questionnaire to assign the refugees to municipalities and to start looking for housing opportunities using the information about family size to find housing of a suitable size.[17] Information about ethnicity was used to create ethnic clusters at the municipality level, which was believed to ease integration.

Importantly for our research design, the allocation decision was based on the questionnaire alone

[16]The questionnaire did not involve any questions on personal characteristics such as education, prior job experience or health.

[17]In practice, the distribution of refugees was carried out in three steps: First, refugees were distributed proportionally to the number of inhabitants in each of the fifteen counties in Denmark. Next, the refugees were allocated to municipalities within counties proportionally to the number of inhabitants in each municipality. In a third and final step the DRC found permanent housing for the resettled refugee within the assigned municipality.

and did not involve any personal meeting between the allocation unit and the refugee prior to allocation. Once allocated to a municipality, the housing officers in the DRC used the questionnaire to look for housing opportunities. Effectively, this meant that the DRC resettled refugees independently of other individual characteristics, and the policy design therefore creates random variation in refugees' initial housing location, conditional on the questionnaire observables. This means that we can compare health outcomes for individuals, who on questionnaire observables were similar, but were allocated to neighborhoods with different income levels to estimate the impact on health of neighborhood income.

The practical implementation of the Spatial Dispersal Policy was influenced by a simultaneous housing shortage.[18] Specifically, the DRC struggled to find enough affordable housing of a suitable size, considering the relatively low income levels of the newly arrived refugees.[19] This shortage is best illustrated by waiting times for permanent housing which were on average six months, but could be up to two years.[20] The effort needed to find permanent housing options is also illustrated by the DRC's need to employ special housing officers (distinct from the refugee's case-worker) who worked full-time on finding housing. The housing shortage implied that the DRC's demand for permanent housing always exceeded the available housing options, and thus effectively created queues of individuals with the same questionnaire observables waiting for permanent housing. This meant that whenever the DRC found a permanent housing opportunity, the DRC offered it to the next refugee in line whom it matched in terms of questionnaire observables. This prevented the DRC from placing refugees in a selective manner.

2.2 Identification

We argue that the design of the Spatial Dispersal Policy made the allocation of individuals random across housing options, conditional on the observables from the questionnaire. This provides us with the variation used for identification. Previous studies have exploited the same natural experiment, arguing that the allocation of refugees was random across municipalities (Damm and Dustmann (2014)) and at the clustered hectare level (Damm (2014)). Our main definition of a neighborhood, a parish, lies somewhere in between these two in terms of the geographic area it spans. In our analysis we will also consider smaller geographical units (stairways of apartment buildings) and municipalities.

For our identification strategy to be valid, we must rule out selection of individuals across neighborhoods. We expect selection of individuals based on the questionnaire observables across neighborhood

[18] See Danish Refugee Council (1991) and Danish Refugee Council (1996).

[19] The DRC was not allowed to buy real estate and rent it to refugees and thus relied solely on renting opportunities.

[20] See Damm (2004) for numbers on waiting times. While waiting for the DRC to find permanent housing, the refugee moved to temporary housing in the municipality that he/she was assigned to within approximately ten days of being granted asylum, see Damm and Dustmann (2014).

types, because the DRC allocated individuals based on these observables. But, once we take this selection into account, we assume that there was no selection into top-, middle- or bottom-income neighborhoods based on other criteria such as individuals' health or educational attainment at arrival which were not included in the questionnaire – i.e. that the income level of the allocation neighborhood was independent of the refugee's individual characteristics not observed by the DRC. We do not assume that the number of individuals allocated to a certain parish or stairway was random, since the supply of affordable housing likely varied across neighborhood income types.

This means that we assume that two individuals who were of similar age, gender, ethnicity and family size were equally likely to find housing in a low-, middle- or top- income municipality – independent of any other potential differences between them. This conditionally random allocation of individuals between municipalities is important even when we let parishes define a neighborhood, because it allows us to compare health outcomes of individuals assigned to parishes in different municipalities.

In a similar way, we assume that once allocated to a municipality, two individuals with the same questionnaire observables had the same probability of finding housing in a low-, middle- or top-income parish independent of any other (un)observable characteristics. We make a completely parallel assumption for selection into stairways of apartment buildings. We argue that these assumptions are valid because individuals were assigned to permanent housing solely based on the questionnaire.

Three concerns arise in this context which could invalidate the design: *i)* the DRC selectively allocated certain types of individuals to certain types of neighborhoods, *ii)* neighborhoods tried to select refugees trough lobbying against/for specific individuals, *iii)* individuals self-selected into neighborhoods. Below, we address each of these concerns carefully. We will address these concerns with a parish in mind as this is the neighborhood level we use throughout most of our specifications. However, a much similar line of reasoning goes for stairways and municipalities.

The scope for the DRC to place individuals in a selective manner was very limited since the housing officer searched for housing based on information from the questionnaire already before the person moved into the municipality. Furthermore, the contemporaneous shortage of housing meant that whenever the DRC found a housing opportunity, there was always a queue of individuals with similar observables waiting for the same type of housing. Therefore the housing option was simply offered to the next person in line. Interviewing the former DRC head of housing, she found it very unlikely that housing officers were able to selectively allocate individuals across neighborhoods due to the constant lack of affordable yet large enough housing options in the housing market.[21] Thus, it seems unlikely

[21] Interview with Bente Bondebjerg on October 22, 2019.

that the DRC systematically placed specific types of individuals in certain types of neighborhoods.

A second concern is that neighborhoods, e.g. through lobbying, tried to affect which types of refugees were allocated to that area. This is a potential issue at all neighborhood levels. At the municipality level the scope for selection was limited due to the short time frame (approximately ten days) from asylum was granted until resettlement took place in the municipality. Once allocated to a municipality, the different parishes could also lobby against/for certain refugees. However, contrary to the municipality, the parishes or stairways did not have a formal administrative unit to organize such lobbying and it therefore seems unlikely that it took place.

Finally, one could worry that the individuals somehow managed to self-select into specific types of neighborhoods. We do not directly observe the actual housing offers made to the refugees but only their first address. It is therefore crucial for our identification strategy that the acceptance rate of housing offers was high. In the previously mentioned interview with the former housing officer, she could not recall that refugees declined a housing offer. The explanation for this is threefold. First, the person only received one housing offer, and if the individual declined that offer, he/she had to move out from the temporary accommodation. This means that there was no bargaining over housing offers and that the cost of declining the offer was high. Second, following the acceptance of a housing offer, the refugee was free to move whenever he/she wanted to. Finally, the difficulty of finding affordable housing was probably even greater for refugees themselves, since they would mostly be without network connections and lack knowledge of the Danish housing market in general. Damm (2009) shows that the take up rate was above 90 percent, which is remarkably high compared to the *Moving to Opportunity* experiment in which the acceptance rate was between 48 and 62 percent.[22]

2.3 Balancing tests

To further support our identifying assumptions, we run a set of balancing tests of neighborhood characteristics on several individual characteristics which were not observed by the DRC housing officer at the time of assignment, but are available to us in the administrative data. At the time of allocation the DRC did not know the educational level and health status of the refugees, which should therefore not correlate with any characteristics of the neighborhood they were assigned to. Thus, to test whether the individuals were distributed randomly across neighborhoods we regress several neighborhood characteristics on the individual refugee characteristics known and unknown to the DRC at the time of allocation. We run the

[22] See Katz, Kling, and Liebman (2001) for numbers on the take up rates in the *Moving to Opportunity* experiment.

following linear regressions:

$$\begin{aligned} y_{n,t-1} &= \alpha + \beta_1 unknown\_educ_{it} + \beta_2 basic\_educ_{it} + \beta_3 academic\_educ_{it} \\ &+ \beta_4 circulatory\_disease_{it} + \beta_5 nutritional\_disease_{it} + \beta_6 neurotic\_disorder_{it} \\ &+ X_{it}\gamma + T_t + \varepsilon_{it} \end{aligned} \tag{1}$$

The neighborhood characteristics, $y_{n,t-1}$, are indicator variables for the poorest, middle or richest third of neighborhoods, and the share of residents suffering from a lifestyle related disease. X_{it} summarizes the individual characteristics known from the questionnaire: age, country of origin, gender, marital status and family size at immigration, and T_t are year of arrival FE.[23]

Table 3 presents the results from these balancing tests. They show that refugees' educational attainments acquired prior to immigration have no significant prediction power of the neighborhood income level or neighbors' health conditions in the initial placement neighborhood.[24] If we use health diagnoses in the first year upon arrival as proxies for refugees' initial health conditions, we find no significant association between initial health and neighborhood income level or neighborhood health.[25] None of the estimated coefficients are statistically different from zero at conventional significance levels, and an F-test of joint insignificance of the education and initial health variables cannot reject that they are jointly equal to zero, see Table 3. Furthermore, we find no evidence of selection on health and education across stairways of apartment buildings or municipalities using similar regression tests.[26]

Based on the balancing tests and the arguments posed in Section 2.2, we argue that the initial neighborhood placement was quasi-random and that we can rule out selection across neighborhoods. The balancing tests underline the importance of conditioning on observables available from the questionnaire. They show that larger families and women were more likely to be assigned to richer neighborhoods. This could be a result of larger families being assigned to cities, in which income was generally higher, and where it was easier to find bigger yet affordable apartments.

[23]We refer to Section 3 for the definition of the neighborhood income groups.

[24]Neighbors' health conditions in the placement parish is measured as the share of residents diagnosed with a lifestyle related disease in the year of a refugee's arrival (yearly incidences).

[25]Unfortunately, we have no ex ante data on refugees' health. However, we do not expect neighborhood quality to have an immediate impact on health. Instead, we expect lifestyle related diseases to build up gradually over time. Thus any difference in the risk of suffering from a lifestyle related disease must be attributed to pre-existing health conditions. One drawback of this measure, is that the detection risk may depend on neighborhood of assignment. One could worry that the detection risk is lower in the low-income neighborhoods.

[26]See Appendix Tables A.1 and A.2. Note that one coefficient is significant at the 5 percent level for the association between municipality level median income and refugees reporting that their highest completed education was basic schooling. This may reflect an imbalance in how refugees' educational attainment was surveyed across municipalities (the survey took place at Danish language training facilities), or it may simply arise by chance, because we are testing multiple hypotheses.

3 Empirical model

The main question posed in this paper is how living in a low-income neighborhood impacts health outcomes. To answer this question we divide all neighborhoods into three equally sized income groups based on their median disposable household income: Bottom-, middle- and top-income neighborhoods. We calculate these groups for each year in our sample and assign all neighborhoods to one of the three groups – regardless of whether the DRC found housing for any individual in a given neighborhood in a given year.

We can use the natural experiment described in Section 2 for identification of causal neighborhood effects in both a reduced form approach and in an instrumental variables (IV) setup. In the reduced form approach we simply estimate the health effects of assignment to a neighborhood of a certain type using OLS. In the IV setup we use the exact same conditionally random variation in assignment neighborhood to instrument the number of years the individual spent in the poorest neighborhoods using 2SLS.[27] This allows us to estimate the average impact of spending one additional year in a low-income neighborhood.

Reduced form strategy. Concretely, in the reduced form setup we estimate the impact on an individual's health outcome $y_{i,t+r}$:

$$y_{i,t+r} = \alpha + \sum_{k=2}^{3} \beta_k \cdot \mathbb{1}[incomegroup_{n,t-1} = k] + \boldsymbol{X_{it}\gamma} + \boldsymbol{T_t} + \boldsymbol{P_n} + \varepsilon_{i,t+r} \tag{2}$$

In model (2), $y_{i,t+r}$ denotes the health outcome of individual i, r years after arrival year t placed in neighborhood n. $incomegroup_{n,t-1}$ denotes the income group of the assignment neighborhood one year prior to arrival $t-1$. We control for the information available from the questionnaire to the DRC: age, country of origin, gender, marital status and family size at immigration summarized in $X_{i,t}$. We also include year of arrival fixed effects, T_t, and parish type fixed effects, P_n. The parish types fixed effects are indicators for urban areas close to big cities, urban areas away from big cities, rural areas close to big cities and rural areas away from big cities.

The coefficients β_k denote the increased risk of diagnosis y if assigned to a middle- or top-income neighborhood relative to being assigned to the poorest neighborhoods. Thus, a negative estimate of β_2 and β_3 means that the risk of being diagnosed with y is lower in a top- and middle-income neighborhood than in a low-income neighborhood. The parameters identify the causal impact of being assigned to a

[27] This is similar in spirit to Angrist and Krueger (1991) who use quarter of birth as an instrument for years of schooling, or Acemoglu and Angrist (2000) who instrument years of schooling using state of birth.

certain type of neighborhood if the allocation of individual i to neighborhood n is random, conditional on the set of included individual characteristics and fixed effects. As we argue in Section 2.2, this assumption of independence is satisfied, since the Spatial Dispersal Policy allows us to rule out selection of individuals into specific neighborhoods if we condition on observables from the questionnaire guiding the allocation.

On top of that, to be sure that the estimated long-term health effect is a result of neighborhood income level, we must rule out individual income effects. For example, if we observe that individuals, who were initially placed in neighborhoods with higher income, have better health outcomes 19 years after immigration, and these individuals at the same time experienced higher income growth, we do not know whether to attribute the improved health outcomes to neighborhood or individual income changes. We test this and provide evidence of the absence of any individual income effects in Section 6.1.

Instrumental variables strategy. We can also use the natural experiment to quantify the health impact of spending one additional year in the poorest neighborhoods. The Spatial Dispersal Policy did not prevent individuals from moving once allocated to a neighborhood, and the decision to relocate most likely depends on individual (unobserved) characteristics along with the amenities of neighborhoods. Therefore, we instrument the number of years spent in the poorest neighborhoods with assignment neighborhood type. This approach yields an estimate which is a weighted average of a series of local average treatment effects (LATE) of one additional year spent in the poorest third of neighborhoods.[28]

A discussion of the assumptions behind our IV strategy is warranted. The Spatial Dispersal Policy provides us with quasi-random variation in initial neighborhoods, conditional on observables, such that the independence assumption is satisfied.[29] Moreover, the initial placement only affected health outcomes through the number of years an individual lived in a specific type of neighborhood, which implies that we can comfortably assume that the exclusion restriction holds.[30] Finally, the income group of the placement neighborhood is a relevant instrument if there is persistence in the type of neighborhood the individual lives in over time. In other words, if the number of years the individual is exposed to a bottom income neighborhood depends on the placement neighborhood income type, our instrument is relevant and has prediction power in the first stage regression.[31] Lastly, we assume monotonicity –

[28] We refer to Angrist and Pischke (2008) for the interpretation of LATE in the case of a multivalued endogenous regressor.

[29] This is discussed in detail in Section 2.2.

[30] As already discussed, in Section 6.1 we show that the initial allocation of individuals did not impact their labor market outcomes. Table 9 shows very precise null-effects on employment and earnings.

[31] The instrumental variables estimation approach also handles if the individual does not move, but the neighborhood changes its rank in the income distribution over time – for example, if an individual is initially placed in a bottom income neighborhood and that neighborhood evolves into a middle income neighborhood over time. We document that the income rank of neighborhoods is highly stable over time (see Appendix Figure A.2).

i.e. that being placed in a bottom income neighborhood always increases years of exposure to bottom income neighborhoods.

These assumptions allow us to scale the neighborhood effects in health by the number of years spent in the poorest third of neighborhoods. We implement the strategy by estimating the following equations with 2SLS:

$$\begin{aligned}
&First\ stage: \quad x_{i,t+r} = \alpha_1 + \sum_{k=1}^{2} \tilde{\beta}_k \cdot \mathbb{1}[incomegroup_{n,t-1} = k] + \boldsymbol{X_{it}\gamma_1} + \boldsymbol{T_t} + \boldsymbol{P_n} + \tilde{\varepsilon}_{i,t+r} \\
&Second\ stage: \quad y_{i,t+r} = \alpha_1 + \beta_k \cdot \hat{x}_{i,t+r} + \boldsymbol{X_{it}\gamma_1} + \boldsymbol{T_t} + \boldsymbol{P_n} + \varepsilon_{i,t+r}
\end{aligned} \tag{3}$$

The predicted number of years an individual has spent r years after immigration t in a bottom income neighborhood is denoted by $\hat{x}_{i,t+r}$, and the controls $\boldsymbol{X_{it}}$, $\boldsymbol{T_t}$ and $\boldsymbol{P_n}$ are the same as in equation (2). Thus, β_k denotes the increased risk of being diagnosed with y following one additional year of exposure to the poorest third of neighborhoods.

4 Data

Our analysis is based on rich administrative data from Statistics Denmark which allows us to link individual records from several registers and track individuals over time. We define our main outcomes of analysis using The National Patient Registry ("LPR"), The Integrated Database for Labor Market Research ("IDA") as well as the Income Register ("IND"). We supplement these longitudinal data sets with the Population Register ("BEF") and information on country of emigration and date of settlement in a Danish municipality from the Migration Register ("VNDS"). Combining these data sets provides us with key demographic variables such as age, gender, origin country and address, and it allows us to identify both relatives and neighbors.

In order to study individuals subject to the Refugee Spatial Dispersal Policy we consider a sample of refugees who arrived between 1986 and 1998. The Migration Register does not carry information on the type of residence permit granted to immigrants in this time period. Instead we define a refugee as someone who emigrated from one of nine refugee sending countries: Afghanistan, Ethiopia, Iran, Iraq, Lebanon, Palestine[32], Sri Lanka and Vietnam in 1986 to 1998, and Somalia 1989 to 1998.[33] We exclude individuals who were married to a Dane or a non-refugee immigrant spouse along with refugees married

[32] Stateless refugees.

[33] See Dustmann, Vasiljeva, and Damm (2018), Foged and Peri (2015), Damm and Dustmann (2014) or Damm (2009) among others for a similar approach. Yugoslavia was also considered a refugee sending country in that time period, but due to the large influx of this particular group the Danish government designed a special dispersal policy for them.

to a refugee spouse arriving more than a year earlier. This prevents the inclusion of individuals who arrived to Denmark as a result of family-reunification – individuals we do not want to include, since they would be living with their spouse instead of being allocated to a municipality through the dispersal policy. Furthermore, we restrict the sample to those aged 18-64 at arrival.

These steps leave us with a sample of 25,738 refugees whose average age at arrival is 30 years. 40 percent of them are female while more than half are married (61 percent). The average family size is 2.4, since many arrive with children, and the two largest ethnic groups in our sample are Iraqi and Lebanese nationals, followed by persons from Somalia and Iran. Upon arrival 30 percent of the sample were surveyed by a statistical agency about their educational level obtained abroad.[34] Of those, 56 percent report basic schooling or less, 21 percent have vocational education while 23 percent arrive with an academic education, c.f. Table 1.

Our main outcomes in the empirical analysis are diagnoses from hospitals based on the National Patient Registry. The diagnoses follow the International Classification of Diseases (ICD) from World Health Organization which carry an extreme level of detail.[35] First, we aggregate the diagnoses, we include in our analysis, into two main groups: lifestyle related diseases and mental disorders. The lifestyle related diseases consist of circulatory diseases[36], nutritional/endocrine/metabolic (referred to as nutritional) diseases[37], chronic obstructive pulmonary disease (COPD), hip arthrosis and alcohol related diseases. The lifestyle related diseases, we include, are the most common lifestyle related diseases (Patienthåndbogen (2017)), and they account for a large share of deaths worldwide (World Health Organization (2018)). The mental disorders considered in our analysis are disorders due to psychoactive substance use, schizophrenic disorders and neurotic disorders.[38]

We study neighborhood effects in lifestyle related diseases because the risk of developing lifestyle related diseases is influenced by individual behavior. That means, that if we expect neighborhoods to influence individual behavior by altering diet or exercise habits, then we would also expect neighborhoods to affect the risk of developing these diseases. Neighborhoods could influence these behaviors for example through the availability of healthy grocery stores or recreational areas but also through the

[34]The information was used for national statistics purposes in an anonymized format, and it was not collected by the DRC.

[35]ICD-8 structure prior to 1994 and thereafter the ICD-10 structure.

[36]Hypertension, ischaemic heart diseases, pulmonary diseases, other forms of heart disease, cerebrovascular diseases and arterial diseases.

[37]Diabetes, obesity and elevated cholesterol levels.

[38]More specifically, we study mental and behavioral disorders due to psychoactive substance use, schizophrenia, schizotypal and delusional disorders, mood (affective) disorders, neurotic, stress-related and somatoform disorders, behavioral syndromes associated with physiological disturbances and physical factors, and disorders of adult personality and behavior. See appendix Section A.1 for a full overview of the grouping of diagnoses.

behavior, attitudes, and appearances of other inhabitants.[39]

Our health measure has the advantage of being very detailed and available for the full population, since health care is universal and provided free of charge to Danish residents, including refugees. However, we do expect under-detection of diseases because not every disease is diagnosed or requires a visit to a hospital.[40]

Second, we study several labor market outcomes to analyze whether our estimated health effects are a result of differences in employment probabilities, earnings or types of occupations across neighborhoods. We measure employment as the fraction of a full working year. This measure takes the value one if the worker was a full-time employee during the whole year. The fraction is less than one and measures the share of a full-time equivalent if the individual was either a part-time employee or not employed in some periods throughout the year. As a measure of income we use information on the annual gross earnings deflated using the consumer price index from Statistics Denmark (with year 2000 as base year) and converted to USD using the exchange rate from the Danish Central Bank on March 27, 2020. In order to characterize occupations according to their task content we use the ratio of communication and cognitive tasks relative to manual tasks in a job.[42] We measure the task content of occupations for those who were employed at the end of November each year.

As previously described, we define a neighborhood as a parish in our baseline specifications, and we will use both phrases interchangeably. For historical reasons, a parish revolves around a church and thus describes smaller neighborhood entities quite well. The individuals in our sample were assigned to 1,055 different parishes which had on average 3,126 inhabitants during the period. We study the importance of small local areas by varying the neighborhood level using a more aggregate level (municipality) and a very fine level considering households living in the same building (stairway level). A parish is a subset of a municipality, whereas a stairway is a subset of a parish. During the period we study, refugees in our sample were distributed across 255 different municipalities and 9,405 different stairways. On average, disregarding the refugees, the municipalities had 15,424 inhabitants whereas a stairway only had 12 inhabitants during the period. For each year we characterize the geographical areas by the median level of household disposable income.[43] The neighborhood income characteristics are supplemented with

[39]See Christakis and Fowler (2007) for examples on how the risk of obesity can be influenced by obese social contacts or Sanbonmatsu et al. (2011) for an overview of how neighborhoods may influence both mental and physical health.

[40]Even though patients can be diagnosed with multiple (and less severe) conditions when visiting the hospital. The detection rate may depend on neighborhood income level since correlational evidence suggests that inhabitants in low-income areas generally utilize health services to a lesser extent than their more affluent counterparts.[41] This may create a downward bias in our estimates.

[42]The task content is from the O*NET database (US Bureau of Labor Statistics) merged to Danish register data using the International Standard Classification of Occupation.

[43]Deflated by the consumer price index (2000 level).

additional neighborhood variables such as the number of GP's per capita, the number of co-nationals, urban/rural area, health care utilization and incidences of lifestyle related diseases and mental disorders among the non-refugee residents. All these characteristics are defined in the same way as individual refugee characteristics, and they are measured one year prior to arrival of each refugee. We refer to Table 2, Table A.8 and Table A.9 for the summary statistics of neighborhood characteristics.

5 Main results

In this section we present our main findings on neighborhood effects in health. We start by presenting the neighborhood effects from the reduced form approach, including evidence showing that these effects differ across age groups and gender. We then proceed to present evidence on the health impacts of spending an additional year in a low-income neighborhood using an IV strategy.

5.1 Reduced form approach

Allocation to the poorest third of neighborhoods increases the risk of developing a lifestyle related disease before 2018 by 1.9 percentage points relative to allocation to the richest third of neighborhoods. The risk of developing a lifestyle related disease is also 1.9 percentage points higher if the individual was allocated to the poorest third of neighborhoods compared to allocation to a middle-income neighborhood, see Panel a of Table 4. This amounts to a 5.4 percent increase in risk relative to the sample mean. These effects are driven by increases in the risk of developing diabetes and hypertensive diseases. Diabetes and hypertensive diseases are subgroups of nutritional and circulatory diseases, which are some of the most common lifestyle related diseases. We do not observe any significant differences in average mental health outcomes across neighborhood income types.

Figure 1 shows that the effect emerges slowly which is consistent with lifestyle related diseases gradually developing over time as a result of health behaviors. Furthermore, the individuals are relatively young at arrival (the mean is 30 years old) and the risk of developing lifestyle related diseases generally increases with age. Most of the effects on health arise 8-15 years after immigration, which is why we focus on this time horizon in Panel b of Table 4. This shows that the risk of developing a lifestyle related disease increases by 1.6 and 1.8 percentage points following allocation to the poorest third of neighborhoods relative to a middle- or top-income neighborhood, respectively.

It is natural to ask whether the increased risk of suffering from a lifestyle related disease in low-income neighborhoods translates into higher mortality rates. We find that individuals placed in low-income neighborhoods have a higher mortality rate than those placed in middle- or top-income neighborhoods, but the difference is small in magnitude and not statistically significant, see Appendix Table A.3.

Our findings in Table 4 are very robust to the choices made in the baseline specification. We find similar results using the mean income instead of median neighborhood income. Using a continuous income measure instead of income group dummies shows that a one standard deviation decrease in median neighborhood income causes an increase in the risk of suffering from a lifestyle related disease of 0.008 percentage points. Finally, we show that the effects are not an artifact of the linear probability model; a probit regression yields the same qualitative effect. As a placebo test, we study some health outcomes that should not be affected by neighborhood income, namely congenital disorders. These tests reveal precise null-effects, confirming that the significant impact on lifestyle related diseases does not simply seem to arise by chance. The full set of robustness checks and placebo tests can be found in Table 5.

5.1.1 Heterogeneous effects

The impact on health of placement neighborhood income type varies significantly by gender and age. Table 6 shows that females experience a larger increase in the risk of developing lifestyle related diseases and nutritional disorders compared to males if they are placed in the poorest third of neighborhoods as opposed to placement in a middle- or top-income neighborhood. In other words, female health is more adversely affected by living in the poorest neighborhoods. More than 19 years after immigration, women placed in the poorest neighborhoods have a 2.9 percentage points higher risk of developing a nutritional disease than men placed in similar neighborhoods, relative to placement in the richest third of neighborhoods.[44] Since we do not observe any differential impacts on diabetes for women, this difference is primarily driven by obesity which is the other large component of the nutritional diagnoses. In our sample a larger share of women than men are diagnosed with nutritional or lifestyle related diseases before 2018, and our estimations indicate that the larger neighborhood effects for females might contribute to this difference. One potential explanation for the differential impact by gender could be that women are more affected by their immediate local environment, because they have lower rates of labor force participation and spend more time at home compared to men.

Another interesting dimension when studying heterogeneous impacts is age, since physical health

[44]The estimate is marginally significant over the full time period, but statistically significant within 8-15 years after immigration.

generally worsens with age. Our findings suggest that placement in the poorest neighborhoods is particularly detrimental for those aged 18-25 at immigration. Their risk of developing a lifestyle related disease prior to 2018 is 3.4 percentage points higher than for older persons placed in similar neighborhoods. This improvement is only observed after more than 15 years after arrival.

If neighborhood income can be considered a proxy for neighborhood quality, our findings on age differences in neighborhood effects imply that the largest improvements in physical health can be reaped by improving neighborhood quality for the younger age groups. This corresponds to the findings in the *Moving to Opportunity* literature which documents larger gains of improving neighborhood quality for younger individuals and children (see Kling, Liebman, and Katz (2007)).

5.2 Instrumental variables approach

In this subsection we turn to the results from the IV approach. First, we learn from Figure 2 that there is substantial persistence in the type of neighborhoods that people live in. After 19 years, those placed in the poorest third of neighborhoods have spent almost 10 years, on average, in that type of neighborhood (Panel a). The behavior for those placed in a middle- or top-income neighborhood is similar, although slightly less persistent (Panels b to c). Furthermore, the graphs reveal that the individuals placed in the poorest neighborhoods have spent significantly more time in a bottom income neighborhood than those placed in a middle or top income neighborhood.[45] This implies that we have a relevant instrument and a very strong first stage (see Table 8).

When we instrument total exposure to each of the three neighborhood income groups, we find that each additional year spent in the lowest income neighborhoods increases the risk of suffering from a lifestyle related disease by 0.3 percentage points. The effects are mainly driven by the occurrence of diabetes and hypertension, see Table 8.[46] The findings are qualitatively similar if we instead instrument average income in all neighborhoods that the individual lived in (See Appendix Table A.5 and Section A.3 for a description of this approach). It is important to instrument the number of years an individual has spent in the poorest neighborhoods, because there is a significant self-selection of less healthy individuals into poorer neighborhoods after the initial allocation. Table A.6 shows that the income gradient in health is larger if endogenous moving is not taken into account.

[45] Appendix Figure A.2 shows that there is substantial persistence in the ranking of neighborhoods in the income distribution.

[46] See Appendix Table A.3 for the dynamics of diagnosed lifestyle related diseases.

6 Mechanisms behind the neighborhood effects

Next, we investigate some of the potential explanations behind the documented neighborhood income gradient in health using the reduced form setup. First, we explore how allocation to a given type of neighborhood affects different individual outcomes that in turn might affect their health outcomes. Second, we investigate how the observed income gradient in health depends on the neighborhoods' characteristics and the composition of residents. Each refugee outcome considered in the first approach and each control variable included in the second approach is testing a different potential explanation, and they capture some of the most obvious (yet measurable) ways in which neighborhoods may affect residents' health outcomes. Taken together, all these tests allow us to rule out a number of plausible explanations behind the negative neighborhood income gradient in health. Finally, we turn to our IV approach to examine the importance of the very local environment and immediate neighbors by varying the size of the neighborhood. We conclude the section by discussing other potential mechanisms which we are not able to measure.

6.1 Individual outcomes

We consider how initial neighborhood allocation affects the individuals' performance in the labor market and their educational attainments after immigration. Differential changes in these outcomes across neighborhoods could potentially contribute to the differences in health outcomes. For example, improved labor market opportunities for individuals in high income neighborhoods could potentially affect health by increasing their life satisfaction and/or by increasing the individuals' income levels.

Labor market. Interestingly, persons allocated to the poorest third of neighborhoods by the Spatial Dispersal Policy do not experience different labor market outcomes than those allocated to top- or middle-income neighborhoods, see Table 9. This implies that the differences in health outcomes are not driven by differential labor market outcomes as a result of initial placement. We estimate very precise zero effects on different measures of employment and income: After more than 19 years in Denmark the cumulative difference in the number of years with any employment is between -0.07 and 0.01 years across the different types of neighborhoods, and it is not statistically significant.[47] Similarly for earnings, we observe differences of less than a typical monthly salary in the cumulative income over 19 years across neighborhoods. This is consistent with the findings in Damm (2014) who documents that living

[47] In general the group of refugees have very weak labor market attachment. The average number of years with any employment during the period considered is 4.19 years.

in socially deprived neighborhoods does not impact the labor market outcomes of refugee men. It is also in line with evidence from the *Moving to Opportunity* experiment. See for example Katz, Kling, and Liebman (2001), Kling, Liebman, and Katz (2007), Sanbonmatsu et al. (2011) or Ludwig et al. (2012) who find no effects on employment, earnings or welfare receipt probability. Thus, we can rule out any income effects of being placed in a bottom, medium or top income neighborhood.

Education. We document a significant difference in educational outcomes across placement neighborhoods. Panel a of Table 10 shows that being placed in a top- or middle-income neighborhood increases the probability of completing an education in Denmark by 1.6 and 2.2 percentage points, respectively, compared to those placed in the poorest third of neighborhoods.[48] The table also shows that these results are primarily driven by completion of vocational education. The combination of Panel a and b shows that the differences in educational attainment across neighborhoods occur within the first eight years after arrival, which is before the observed differences in health outcomes across neighborhoods arise.

It cannot directly be inferred from Table 10 whether the increased educational level decreases the risk of developing lifestyle related diseases. More education might lead to higher employment probabilities and also higher wages which in turn might affect health directly and indirectly. However, Table 9 shows that the increased educational level among individuals placed in richer neighborhoods does not translate into more employment or higher earnings, on average. Second, increased educational levels may increase knowledge about health related topics. However, Table 10 shows that the probability of completing a health specific education does not differ across neighborhoods. Third, even though earnings are not affected, better educated individuals may be employed in jobs that are less detrimental to health, for example by finding employment in less physically demanding jobs. Column 5 in Table 9 shows that the occupations where the individuals are employed do not differ in task complexity across neighborhoods.[49] Fourth, more education can increase general knowledge and the ability to follow and understand general health guidelines and advice from health professionals and authorities. Finally, obtaining an education could improve self-esteem or impact the formation of social networks which in turn might improve general well-being, and thus possibly health outcomes in the long term. Based on the timing of completion of education, the two latter explanations may be at play for the population we study.

[48] The results are very similar if we study enrollment instead of completion.

[49] We define occupations by their manual, cognitive and communicative task content. Our results show that there are no significant differences in each of these task contents or a combined index of the three.

6.2 Neighborhood characteristics and residents

Neighborhood characteristics. Turning to the characteristics of the neighborhood, we can rule out that the income gradient in health is driven by differences between urban and rural areas or local institutions at the municipality level (see Table 11). The Danish health care is universal and provided free of charge to all residents, including refugees. This makes it unlikely that the differential health outcomes are driven by differences in access to health care. Moreover, all residents have access to medical treatment of virtually the same quality. However, there might be minor differences in health care access and quality across geographical areas. Residents in rural areas may have restricted access to the health care system because they generally travel longer distances to visit their GP or local hospital. The characteristics of the neighborhood can also differ systematically between rural and urban areas in terms of education possibilities, spare time activities, air pollution etc. However, we find no evidence that such differences between rural and urban areas explain the income gradient.

By including municipality fixed effects we further control for such differences between areas. Comparing neighborhoods within the same municipality allows us to compare neighborhoods that are subject to the same local authorities. Even though hospitals and overall health policy was run by the counties throughout the period, municipalities could still affect access to health care such as rehabilitation offers or health preventive actions. The local authorities might also differ in their tax rate and service level (such as spending per pupil, policemen/inhabitant ratio or cultural investments). Moreover, characteristics of health care professionals may also differ between municipalities but less so within municipalities.[50] We find no evidence that the income gradient can be attributed to differences across municipalities even though our estimates become less precise when including municipality fixed effects, due to the lack of statistical power.

As an alternative to including municipality fixed effects, we include the number of general practitioners per inhabitant in the municipality as a control variable, which supports the conclusion that differences in access to health care does not explain the income gradient.[51]

Neighborhood residents. It is likely that health behavior in the neighborhood could be transmitted to the resettled refugees who may adopt health behaviors from their peer groups. One way to study this is by controlling for the fraction of inhabitants who suffer from a lifestyle related disease in the year

[50] Especially in large municipalities they might also differ within municipalities.

[51] The conclusion remains if we control for the number of general practitioners per capita along with municipality expenditures on social and health services, see Appendix Table A.4. However, municipality expenditures on health services may reflect both the quality of health services and the health conditions of inhabitants.

prior to resettlement.[52] We find that allocation to a neighborhood in which a larger share of residents is diagnosed with a lifestyle related disease significantly increases the risk of developing a lifestyle related disease. However, the estimates presented in Table 11 indicate that this is not a main driver behind our results as the estimated coefficients are very similar to the baseline estimates.

If we instead include the number of co-nationals as a control variable, we reach the same conclusion which suggests that the presence of ethnic networks is not an important factor behind the results.[53]

Lastly, inhabitants with very low income levels may be a relevant peer group for the refugees who also have very low income levels. We therefore include the poverty rate in a neighborhood as a control variable, but it does not have much explanatory power with regards to the income gradient from the baseline results.

6.3 Varying the neighborhood size

Taking one step further, we explore the mechanisms behind the results by varying the neighborhood size. Specifically, if the health outcomes are driven by interaction with peer groups, we would expect effects to become larger in magnitude as the measurement of peer groups becomes more accurate. Thus, measuring median income at the parish level rather than at the municipality level should bring us closer to the income levels of peers as the population becomes smaller and the probability of interaction is increased. The same argument goes for measuring median income levels in the apartment building (more specifically, a particular stairway of an apartment complex) rather than measuring income levels at the parish level.

When we estimate the effect of exposure to the poorest neighborhoods in this way, the estimated coefficients do not change much at the parish level compared to the municipality level (see Table 12). Six of seven coefficients become larger in magnitude when moving from the municipality or parish to the stairway level. Especially one estimate is large and precisely estimated in this case; the impact on diabetes, which consistently seems to be driving impacts on nutritional diagnoses. Moreover, diabetes is one of the major components of lifestyle related diseases, and it is the single most prevalent disease in the refugee population. The analysis across neighborhood levels reveals that the estimated impacts on

[52] We measure this control at the municipality level, since we do not have a good measure for prevalence before 1994 at the parish level, and the number of inhabitants diagnosed in a given year fluctuates relatively much due to low numbers of inhabitants in some parishes.

[53] Ideally, we want to measure the income levels among co-nationals, but this is not feasible because the number of co-nationals prior to immigration is very low in a number of neighborhoods. As an alternative to including the number of co-nationals, we use the number of individuals from refugee sending countries and the share of all immigrants in the neighborhood in Appendix Table A.4. We also include the average income among immigrants in the neighborhood as a control, but this does not affect the estimates much either.

diabetes and lifestyle related diseases are more than tripled when considering median income levels of immediate neighbors compared to the median income of more distinct peers in the parish.[54] For lifestyle related diseases an additional year spent in the poorest stairways increases the risk by 1.2 percentage points compared to 0.4 percentage points for the poorest parishes. This suggests that the characteristics of the very local neighborhood is an important factor for health outcomes. This may be due to a transmission of health behaviors from the immediate neighbors and the exposure to the characteristics of a very small geographical area.

6.4 Remaining explanations

What are the remaining differences between the poorest and richest neighborhoods once we sum up the results from Section 6.2 and Section 6.1? Some of the effects may be due to different educational outcomes for refugees. We can, among other things, rule out both individual income effects and municipality level differences across neighborhoods as well as the presence of ethnic networks as important explanations. This may reflect that what matters most for the health outcomes, we study, are the characteristics of the very local neighborhood such as the characteristics and behaviors of the immediate neighbors, along with the supply of food/grocery store options, immediate recreational areas and local sports clubs. Using the income of the immediate neighbors as a proxy for the very local neighborhood quality, our results from Section 6.3 indicate that such characteristics of the very local environment are important.

Given our results, especially amenities related to diet or exercise or behavior of immediate neighbors could potentially be very important, since both diet and exercise matter for the risk of developing lifestyle related diseases in general, and diabetes in particular. Neighborhood characteristics such as traffic noise or air pollution may be less important determinants of diseases such as diabetes.[55]

Finally, since we do not control for the quality of the apartments that the DRC assigned the individuals to, it is possible that we capture apartment effects in health as opposed to neighborhood effects – i.e. that it is in fact the low quality apartments in the poorest neighborhoods that we measure the effect of. We do not observe the quality of the assigned apartments, but since we can rule out individual income effects, we can rule out large differences in apartment rents, which we in general would expect to be

[54] In an alternative approach we simultaneously instrument the number of years spent in a bottom income parish and a bottom income stairway, respectively. This exercise shows that an additional year spent in a bottom income stairway increases the risk of diabetes by 0.9 percentage points, whereas this estimate is only -0.1 percentage points and not significant for time spent in a low-income parish (see Appendix Table A.7).

[55] Note that our measure of lifestyle related diseases does not include asthma. However, air pollution or traffic noise may be indirectly linked to any disease caused by factors such as stress, happiness etc.

correlated with quality. That is, the apartment quality could only to a limited extent be reflected in prices and still be within the refugees' budget. Yet, the price for quality may vary across the country such that individuals in rural areas far away from the capital got better quality apartments for the same rent as those placed in cities. However, the neighborhood income gradient persists even when we compare individuals placed in the same municipality and control for parish types – i.e. we compare rural parishes with rural parishes in the same municipality. Thus, we do not believe this is the main explanation behind our results.

7 External validity

Even though the persons who were resettled through the Spatial Dispersal Policy are diverse in terms of age, gender, family type, country of origin etc., they may be affected differently by neighborhoods than the general population, because they come from very different circumstances. However, the general population is not randomly allocated across neighborhoods which makes it extremely challenging to identify causal neighborhood effects in health for natives in a similar fashion as can be done for the refugees. Therefore, to study generalizability, we apply a completely naive approach to estimating neighborhood effects in health to verify whether similar effects exist for the native population. In this naive approach, we estimate neighborhood effects in health by comparing health outcomes of non-refugee residents turning 30 between 1994 and 1998, who had the same gross income at age 30 (to rule out ex ante income differences) and lived in the same municipality, but lived in either low-, middle- or top-income neighborhoods.[56] We then estimate if living in a bottom-, middle- or top income parish at age 30 is associated with different risks of suffering from a lifestyle related disease or mental disorder. Thus, we estimate a model similar to model (2) with a slightly different set of control variables.[57]

In general, the estimated coefficients in Table 13 are similar to those presented in Table 4. In the naive approach for the native Danish population we also find a negative neighborhood income gradient in lifestyle related diseases. Living in the poorest third of neighborhoods at age 30 is associated with a significantly higher risk of suffering from any of the lifestyle related diseases and mental disorders

[56] Age 30 corresponds to the mean age for refugees at year of immigration.

[57] The estimated model:

$$y_{i,t+r} = \alpha + \sum_{k=2}^{3} \beta_k \mathbb{1}[incomegroup_{n,t-1} = k] + \boldsymbol{\overline{X}_{it}\gamma} + +\boldsymbol{P_n} + \boldsymbol{M_n} + \varepsilon_{i,t+r} \tag{4}$$

$\overline{X}_{it}$ denotes individual controls (marital status, gender and logarithmic individual gross income at age 30). The large sample size allows us to include municipality fixed effects, $\boldsymbol{M_n}$, in addition to the parish type fixed effects from our baseline estimation, which accounts for any institutional differences across municipalities. Otherwise the variables are similar to those in model (2).

considered for refugees. This is the case both if we consider the risk of developing the disease 8-15 years after the 30^{th} birthday (Panel b) or anytime before 2018 (Panel a). Thus, Table 13 suggests that our main findings do not only apply to refugees.

The neighborhood effect estimates in Table 13 are smaller in magnitude than those in Table 4, but since the share of natives suffering from any of these diseases is also lower than for refugees, the neighborhood effects for natives, measured in percentage terms, are actually larger than among refugees.[58] This is in accordance with our expectations, because self-selection of unhealthy individuals into poorer neighborhoods will create a negative selection bias. Thus, our estimates for natives are a sum of the causal neighborhood effects in health and the selection into neighborhoods, which our approach for natives does not fully account for. For the native population, identification would require that individual (unobserved) characteristics are uncorrelated with the type of neighborhood that the individual lives in at age 30, conditional on the included controls.

In line with this, the negative neighborhood income gradient in mental health presented in Table 13 therefore indicates that the neighborhood income gradient in these disorders is primarily caused by selection, since we fail to document a causal impact on mental health for the average refugee. For the other diseases considered, taken together Tables 4 and 13 indicate that the negative correlation between neighborhood income and individual health is a result of both selection into the poorest neighborhoods as well as a causal effect.

8 Concluding remarks

We study a Spatial Dispersal Policy in act from 1986 to 1998 which quasi-randomly resettled individuals in different neighborhoods. This natural experiment allows us to rule out selection of individuals into neighborhoods and provide causal estimates of the impacts of neighborhoods on residents' health. Specifically, we characterize neighborhoods by their median income levels to study how the risk of developing a number of lifestyle related diseases and mental disorders depends on the income of the neighborhood in which the person was resettled.

We document that individuals who were resettled in the poorest third of neighborhoods are at higher risk of suffering from a lifestyle related disease. We provide evidence that the risk of developing a lifestyle related disease increases with the number of years spent in a low-income neighborhood, and

[58] For instance, there is a 7.0 percent higher risk among natives and a 5.4 percent higher risk among refugees of ever being diagnosed with a lifestyle related disease if living in a low-income neighborhood at age 30 (for natives) or placement in a low-income neighborhood (for refugees).

this is primarily driven by an increased risk of suffering from diabetes and hypertension. Furthermore, we show that exposure to the poorest neighborhoods is particularly harmful for women. On average, mental health is not affected by the neighborhood type.

We contribute to the understanding of neighborhood effects in health by examining a number of potential mechanism that have not been tested previously. The neighborhood income gradient in health cannot be explained by differences in individuals' employment or earnings across neighborhoods, but we document that persons assigned to the richest neighborhoods are more likely to obtain a vocational non-health related education post-immigration. We can rule out that the impacts on health outcomes are caused by differences across municipalities, and we show that the income gradient is not a result of the presence of ethnic networks or differences in poverty rates. Remaining explanations for the observed income gradient include differences in neighborhood amenities and the health behaviors of residents, and we provide evidence that what matters most for neighborhood effects in health is the very local neighborhood. The income level of immediate neighbors living in the same stairway of an apartment building is more important for health outcomes than the income levels of those living in the same parish.

Thus, studying how immediate neighbors' exercise, diet and smoking habits and the access to local recreational areas affect residents' behavior could provide a better understanding of the neighborhood effects in health documented in this paper. Such an understanding can serve as a guideline for policy interventions aimed at improving health conditions in the poorest neighborhoods.

Table 1: Summary Statistics for the Population of Refugees

	All Mean	Bottom Mean	Middle Mean	Top Mean
Characteristics at Immigration				
Age	30.58	30.06	30.96	30.79
Female	0.40	0.40	0.41	0.40
Married	0.61	0.64	0.62	0.61
Number of Family Members	2.36	2.20	2.38	2.45
Number of Children	0.84	0.74	0.85	0.90
Origin Country				
Iraq	0.20	0.24	0.19	0.19
Lebanon	0.20	0.13	0.18	0.22
Somalia	0.18	0.27	0.18	0.15
Iran	0.17	0.11	0.15	0.19
Sri Lanka	0.12	0.13	0.14	0.11
Vietnam	0.08	0.06	0.11	0.08
Afghanistan	0.03	0.04	0.03	0.04
Ethiopia	0.02	0.02	0.01	0.02
Education Surveyed				
Basic Education	0.56	0.54	0.58	0.55
Vocational Education	0.21	0.22	0.20	0.21
Academic Education	0.23	0.25	0.22	0.24
Education Not Surveyed	0.70	0.70	0.70	0.70
N	25,738	4,288	7,654	12,406

Notes: Summary statistics for the full sample of refugees and by parish income groups. The sample consists of refugees between 18-64 years of age who arrived to Denmark between 1986 to 1998 from Iraq, Lebanon, Somalia, Iran, Sri Lanka, Vietnam, Afghanistan and Ethiopia. We do not include arrivals based on family-reunifications. All refugee characteristics are measured at year of immigration. Basic, vocational and academic education is only available for those who were surveyed. Column ”All” presents the mean of characteristics among all refugees in our sample irrespective of parish income group. ”Bottom” refers to characteristics among refugees assigned to the bottom third of parishes measured by median disposable income in a given year. Similarly, ”Middle” and ”Top” refer to characteristics among refugees assigned to the middle and top third of parishes measured by disposable income, respectively. The parish income groups are defined among all parishes, irrespective of any refugee assignment. We define income group of assignment parish one year prior to immigration by median disposable income among all inhabitants aged 18 or above. Data is from administrative registers provided by Statistics Denmark.

Table 2: Summary Statistics for Initial Placement (Parish)

	Bottom Mean	Middle Mean	Top Mean
Characteristics of Residents			
Age	46.48	46.85	45.61
Median Household Income	13,953.39	14,602.77	16,017.42
Employment Rate	0.63	0.68	0.74
Prevalence of Lifestyle Related Diseases	0.09	0.08	0.07
Inhabitants	3,987.00	4,351.20	5,311.90
Co-Nationals	17.49	12.30	8.79
Poverty Rate	0.09	0.07	0.05
Parish Type			
Urban Area (Near City)	0.45	0.43	0.68
Urban Area (Away from City)	0.04	0.19	0.16
Rural Area (Near City)	0.09	0.10	0.08
Rural Area (Away from City)	0.30	0.21	0.05
Characteristics of Municipality			
General Practioners per 1,000 Inhabitants	0.46	0.43	0.46
Incidences of Lifestyle Related Diseases per 1,000 Inhabitants	33.01	29.31	26.11
Health and Social Expenditures per Capita	4,016.16	4,112.72	4,022.29
N	683	1,456	2,773

Notes: Summary statistics for parishes in which refugees were resettled. ”Bottom”, ”Middle” and ”Top” refer to parish characteristics of parishes in the bottom, middle and top third of parishes measured by median parish disposable income in a given year. We calculate the median income of each parish including all inhabitants in each parish aged 18 or above and define the income groups among all parishes, irrespective of any refugee assignment. All parish characteristics are measured one year prior to immigration. Employment rate is the share of the population with any employment between the ages of 18-64. Prevalence of lifestyle related diseases is measured as all incidences over the previous 8 years and thus only defined for refugees arriving after 1993. Health and social expenditures per capita and median household income is measured in USD. Observations are parish-year. Data on ”Health and Social Expenditures per Capita” stems from Statistikbanken, (REG1 and REG11). Parish types are defined by Ministry for Cities, Housing and Rural Areas (2013). All other data is from administrative registers provided by Statistics Denmark.

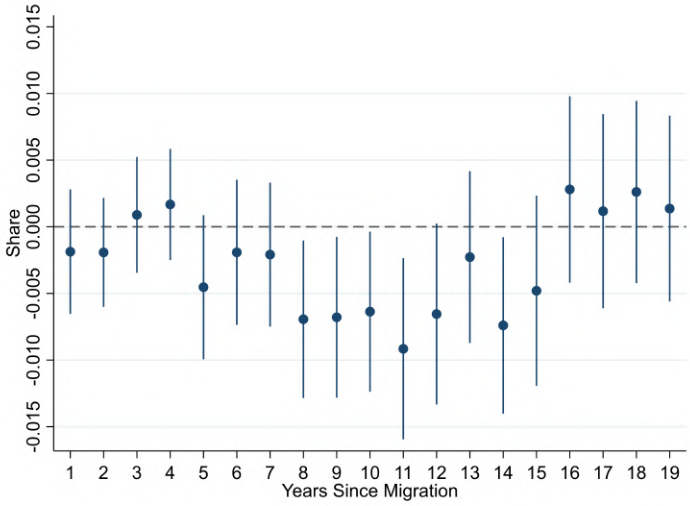
Table 3: Balancing Tests

	(1) Bottom Income Group	(2) Middle Income Group	(3) Top Income Group	(4) Lifestyle Related
Unobserved at Time of Allocation				
Unknown Education	0.000	0.008	-0.008	-0.000
	(0.010)	(0.013)	(0.014)	(0.000)
Basic Education	-0.001	0.024	-0.023	-0.000
	(0.011)	(0.014)	(0.015)	(0.000)
Academic Education	0.011	0.003	-0.014	0.000
	(0.013)	(0.016)	(0.018)	(0.000)
Circulatory Disease	-0.001	-0.027	0.027	0.000
	(0.022)	(0.028)	(0.030)	(0.000)
Nutritional Disease	-0.002	-0.017	0.019	0.001
	(0.031)	(0.038)	(0.041)	(0.001)
Neurotic Disorder	-0.086	0.044	0.042	0.001
	(0.049)	(0.073)	(0.078)	(0.001)
Observed at Time of Allocation				
Age 30-49 Years	-0.003	-0.002	0.005	-0.000
	(0.006)	(0.007)	(0.008)	(0.000)
Age 50-64 Years	-0.022**	0.031**	-0.009	0.000
	(0.009)	(0.013)	(0.014)	(0.000)
Female	-0.003	-0.006	0.010	0.000
	(0.004)	(0.006)	(0.006)	(0.000)
Number of Adults	-0.015	-0.011	0.026**	-0.000***
	(0.010)	(0.009)	(0.011)	(0.000)
Number of Children	-0.002	0.022**	-0.020	-0.000**
0-2 Years Old	(0.009)	(0.011)	(0.012)	(0.000)
Number of Children	-0.007**	0.003	0.004	-0.000***
3-17 Years Old	(0.003)	(0.003)	(0.004)	(0.000)
Married	0.013**	0.002	-0.015	0.000***
	(0.006)	(0.008)	(0.008)	(0.000)
Year of Immigration FE	Yes	Yes	Yes	Yes
Country of Origin FE	Yes	Yes	Yes	Yes
Parish Type FE	No	No	No	No
Municipality FE	No	No	No	No
N	24,348	24,348	24,348	24,484
F	0.74	1.02	0.80	1.33
Pr > F	0.62	0.41	0.57	0.24

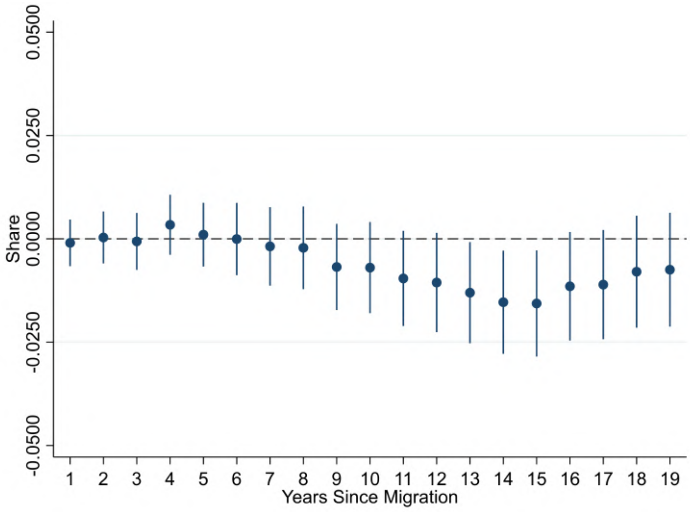
Notes: Balancing tests for parishes using linear regressions. Standard errors in parentheses clustered at the household level. $^{**}p < 0.05, ^{***}p < 0.01$. F denotes the F-statistic for joint insignificance of the educational attainment dummies and pre-existing health conditions. Each column represents a different balancing test testing if refugees with certain characteristics (column farthest to the left) are more likely to be placed in parishes with specific characteristics (dependent variables). The dependent variables in (1)-(3) are dummies for assignment to the bottom third income parish (1), middle third income parish (2) or top third income parish (3). In column (4) the dependent variable is the incidence (as a share of inhabitants) of lifestyle related diseases. The controls are individual characteristics observed by the DRC at time of assignment and characteristics which the DRC does not observe at time of assignment: initial education and initial health. As a proxy for initial health we use diagnoses within the first year upon arrival, but measure all other individual characteristics at year of immigration. We measure all parish characteristics one year prior to immigration.

Figure 1: Development of Lifestyle Related Diagnoses

(a) Yearly Incidents



(b) Cumulative



Notes: Standard errors in parentheses clustered at parish × immigration year level. 90 percent confidence intervals. The graphs plot the development of lifestyle related diseases over time. The coefficients plotted show the increased probability of being diagnosed with lifestyle related diseases if initially assigned to a top-income neighborhood compared to a bottom-income neighborhood. In Panel (a) we show the coefficients from 19 different regression, one for each year plotted, in which the dependent variable is a dummy for being diagnosed with a lifestyle related disease in the year considered. In Panel (b) the coefficients also stem from 19 different regressions but the dependent variable in this panel is a dummy for being diagnosed in the year considered or any year before that since year of immigration. We measure parish income groups one year prior to arrival based on median disposable income in each parish among all parishes in Denmark in a given year.

Table 4: Main Results

	Lifestyle Related	Circulatory	Nutritional	Hypertension	Diabetes	Mental Disorder	Neurotic
	(a) Ever diagnosed						
Middle	-0.019*	-0.025***	-0.013*	-0.018***	-0.009	-0.003	0.001
	(0.010)	(0.009)	(0.008)	(0.007)	(0.007)	(0.010)	(0.008)
Top	-0.019**	-0.017*	-0.016**	-0.020***	-0.014**	-0.004	-0.000
	(0.009)	(0.009)	(0.008)	(0.007)	(0.007)	(0.009)	(0.008)
	(b) Diagnosed 8-15 years after immigration						
Middle	-0.016**	-0.007	-0.012**	-0.005	-0.006	0.006	0.003
	(0.008)	(0.006)	(0.006)	(0.004)	(0.005)	(0.007)	(0.005)
Top	-0.018**	-0.004	-0.016***	-0.003	-0.009*	0.011*	0.006
	(0.007)	(0.006)	(0.006)	(0.004)	(0.005)	(0.006)	(0.005)
Sample Mean	0.35	0.25	0.20	0.12	0.13	0.24	0.16
N	22,948	22,948	22,948	22,948	22,948	22,948	22,948
Parish Type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	No	No	No	No	No	No	No

Notes: Standard errors clustered at parish $\times$ immigration year level. $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$. Estimates from a linear probability model testing the impact of assignment parish income group on the probability of being diagnosed with each of the diseases in the top panel. The estimates show the the increased probability of being diagnosed with each of the considered diseases if assigned to the middle third or top third income neighborhoods compared to a bottom third income neighborhood. In Panel (a) the dependent variable is an indicator for being diagnosed with the disease considered at some point from year of arrival before 2018. In Panel (b) the dependent variable is a dummy for being diagnosed with the disease considered disease 8-15 years after immigration. We measure parish income groups one year prior to arrival based on median disposable in each parish among all parishes in Denmark in a given year. We control for individual characteristics observed at time of assignment by including controls for gender, marital status, family size, and country of origin as well as age and year fixed effects. The sample mean denotes the share of refugees diagnosed with the disease in the top panel before 2018.

Table 5: Robustness Checks and Placebo Tests

	Panel A: Robustness of Lifestyle Related Diseases				Panel B: Placebo Test of Congenital Disorders	
	Baseline	(1)	(2)	(3)	Abnormalities	Metabolic
			(a) Ever diagnosed			
Middle	-0.019*	-0.018*		-0.018*	0.002	0.001
	(0.010)	(0.010)		(0.010)	(0.004)	(0.005)
Top	-0.019**	-0.018*		-0.019**	0.000	-0.002
	(0.009)	(0.010)		(0.010)	(0.004)	(0.005)
Bottom		0.000				
		(.)				
Standardized Median Income			-0.008**			
			(0.004)			
			(b) Diagnosed 8-15 years after immigration			
Middle	-0.016**			-0.017**	-0.002	0.001
	(0.008)			(0.007)	(0.003)	(0.002)
Top	-0.018**	-0.019**		-0.019***	-0.003	0.002
	(0.007)	(0.008)		(0.007)	(0.002)	(0.002)
Middle		-0.019**				
		(0.008)				
Standardized Median Income			-0.008***			
			(0.003)			
N	22,948	22,948	22,948	22,948	22,948	22,948
Parish Type FE	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	No	No	No	No	No	No
Income Type	Disposable	Disposable	Disposable	Disposable	Disposable	Disposable
Moment	Median	Mean	Continuous	Median	Median	Median
Method	OLS	OLS	OLS	Probit	OLS	OLS

Notes: Standard errors in parentheses clustered at parish $\times$ immigration year level. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$. All estimates in Panel A show the impact of assignment parish on the probability of being diagnosed with a lifestyle related disease in different setups. In Panel B we use congenital disorders (congenital abnormalities and congenital metabolic disorders) as placebo outcomes which should not be affected by neighborhood characteristics. Column (Baseline) replicates the main results from Table 4. Column (1) show the same estimation where income groups instead are based the mean parish income. Column (2) demonstrates the estimated effects using a standardized continuous income measure, and column (3) shows the estimated neighborhood effects from a probit model. In Panel (a) the dependent variable is an indicator for being diagnosed with a disease at some point from year of arrival before 2018. In Panel (b) the dependent variable is an indicator for being diagnosed with a disease 8-15 years after immigration. We measure parish characteristics one year prior to arrival. In all regressions we control for individual characteristics observed at time of assignment by including controls for gender, marital status, family size, and country if origin as well as age and year fixed effects.

Table 6: Heterogeneous Effects by Gender

	Lifestyle Related	Circulatory	Nutritional	Hypertension	Diabetes	Mental Disorder	Neurotic
	(a) Ever diagnosed						
Middle × Female	-0.018	0.001	-0.008	-0.010	0.012	-0.011	-0.009
	(0.020)	(0.016)	(0.018)	(0.013)	(0.013)	(0.016)	(0.014)
Top × Female	-0.029	-0.008	-0.029*	-0.023*	-0.003	-0.004	0.004
	(0.019)	(0.015)	(0.017)	(0.012)	(0.012)	(0.015)	(0.013)
	(b) Diagnosed 8-15 years after immigration						
Middle × Female	-0.041**	-0.010	-0.035***	-0.013	0.006	-0.014	0.003
	(0.016)	(0.012)	(0.013)	(0.008)	(0.010)	(0.012)	(0.010)
Top × Female	-0.044***	-0.012	-0.039***	-0.015**	0.002	-0.012	0.005
	(0.016)	(0.011)	(0.013)	(0.008)	(0.009)	(0.012)	(0.009)
N	22,948	22,948	22,948	22,948	22,948	22,948	22,948
Parish Type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	No	No	No	No	No	No	No

Notes: Standard errors in parentheses clustered at parish × immigration year level. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$. The table shows estimates from a linear probability model testing gender differences in the impact of assignment parish income group on the probability of being diagnosed with each of the diseases in the top panel. In panel (a) the dependent variable is a dummy for being diagnosed with a lifestyle related disease at some point from year of arrival before 2018. In panel (b) the dependent variable is a dummy for being diagnosed with a lifestyle related disease 8-15 years after immigration. We measure parish income groups one year prior to arrival based on median disposable in each parish among all parishes in Denmark in a given year. In all regressions we control for individual characteristics observed at time of assignment by including controls for gender, marital status, family size, and country of origin as well as age and year fixed effects.

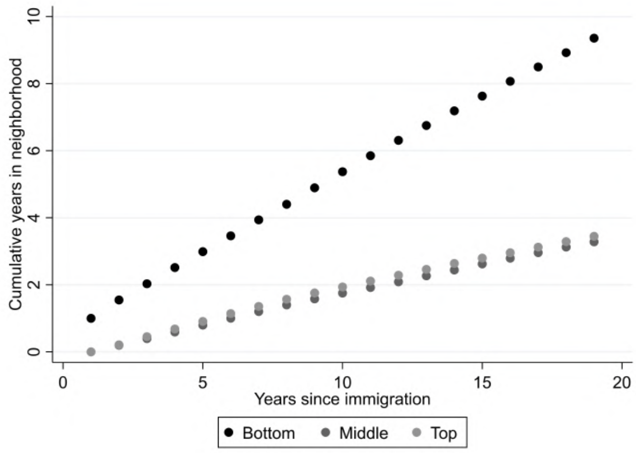
Table 7: Heterogeneous Effects by Age

	Lifestyle Related	Circulatory	Nutritional	Hypertension	Diabetes	Mental Disorder	Neurotic
	(a) Ever diagnosed						
Middle × Age 18-25	-0.009	0.001	-0.015	0.002	-0.003	0.007	0.010
	(0.019)	(0.016)	(0.015)	(0.012)	(0.012)	(0.018)	(0.015)
Top × Age 18-25	-0.034**	-0.018	-0.025*	-0.007	-0.007	-0.018	0.003
	(0.017)	(0.014)	(0.014)	(0.011)	(0.012)	(0.017)	(0.014)
	(b) Diagnosed 8-15 years after immigration						
Middle × Age 18-25	-0.011	0.005	-0.005	0.008	0.012	-0.007	-0.003
	(0.014)	(0.010)	(0.011)	(0.006)	(0.009)	(0.013)	(0.010)
Top × Age 18-25	-0.012	-0.001	-0.004	0.004	0.016*	-0.022*	-0.003
	(0.013)	(0.009)	(0.010)	(0.006)	(0.008)	(0.012)	(0.009)
N	22,948	22,948	22,948	22,948	22,948	22,948	22,948
Parish Type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	No	No	No	No	No	No	No

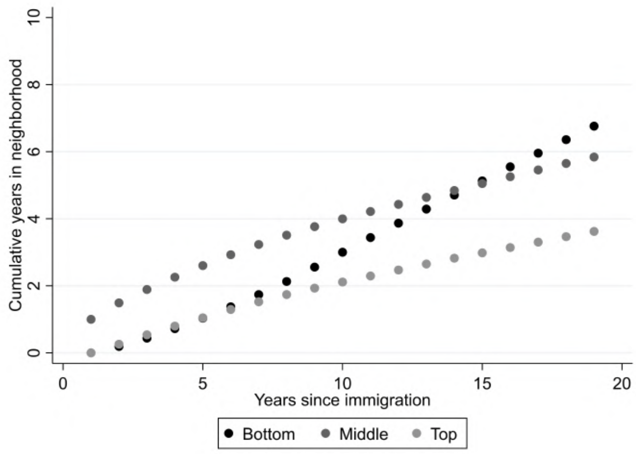
Notes: Standard errors in parentheses clustered at parish × immigration year level. $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$. The table shows estimates from a linear probability model testing age differences in the impact of assignment parish income group on the probability of being diagnosed with each of the diseases in the top panel. In panel (a) the dependent variable is a dummy for being diagnosed with a lifestyle related disease at some point from year of arrival before 2018. In panel (b) the dependent variable is a dummy for being diagnosed with a lifestyle related disease 8-15 years after immigration. We measure parish income groups one year prior to arrival based on median disposable in each parish among all parishes in Denmark in a given year. In all regressions we control for individual characteristics observed at time of assignment by including controls for gender, marital status, family size, and country of origin as well as age and year fixed effects.

Figure 2: Cumulative Exposure to Neighborhoods by Years Since Immigration

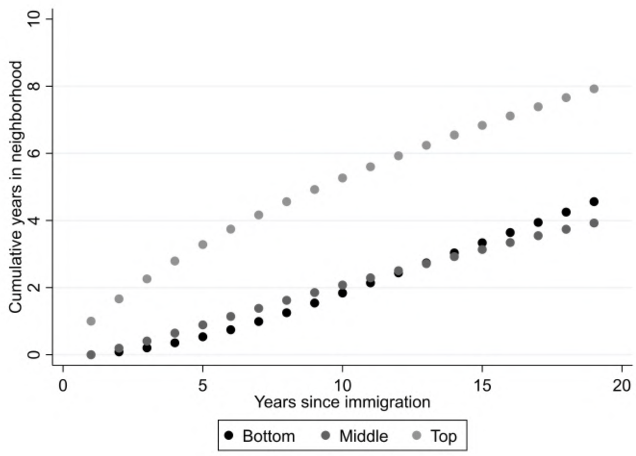
(a) Placed in Bottom Income Neighborhood



(b) Placed in Middle Income Neighborhood



(c) Placed in Top Income Neighborhood



Notes: The figure plots the cumulative exposure to bottom, middle and top third income neighborhoods conditional on type of initial placement neighborhood against years since immigration. Panel (a) shows the cumulative exposure to each neighborhood type among those refugees initially placed in the bottom third income neighborhoods. Similarly, Panel (b) and Panel (c) show the cumulative exposure to the different neighborhood types among those initially placed in the middle third or top third income neighborhoods, respectively. We measure parish income groups one year prior to arrival based on median disposable in each parish among all parishes in Denmark in a given year.

Table 8: IV Estimates

	Lifestyle Related	Circulatory	Nutritional	Hypertension	Diabetes	Mental Disorder	Neurotic
Years of Exposure	0.003*	0.002	0.003*	0.003**	0.002**	0.001	0.000
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
N	22,948	22,948	22,948	22,948	22,948	22,948	22,948
Parish Type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	No	No	No	No	No	No	No

Notes: Standard errors in parentheses clustered at parish $\times$ immigration year level. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$. The table shows the increased risk of being diagnosed with one of the diseases in the top panel following an additional year of exposure to a bottom income bottom neighborhood. We use initial placement neighborhood income group as an instrument in the first stage. The F-statistic from the first stage regression for years of exposure to a bottom income neighborhood is 283.06, and the estimated coefficients from the first stage are $\tilde{\beta}_1 = 5.35$ and $\tilde{\beta}_2 = 2.62$. We measure parish income groups one year prior to arrival based on median disposable in each parish among all parishes in Denmark in a given year. In all regressions we control for individual characteristics observed at time of assignment by including controls for gender, marital status, family size, and country of origin as well as age and year fixed effects.

Table 9: Labor Market Outcomes

	Employment>0	Employment	Labor Income	Business Income	Task Complexity
	(a) Cumulative since immigration				
Middle	-0.07	-0.09	-3,773.02	-3,462.35	-0.00
	(0.12)	(0.10)	(4,303.51)	(4,470.22)	(0.02)
Top	0.01	-0.03	-1,632.94	-1,805.64	-0.02
	(0.11)	(0.10)	(4,121.17)	(4,286.96)	(0.02)
	(b) 8-15 years after immigration				
Middle	-0.06	-0.06	-2,257.41	-1,959.54	-0.02
	(0.06)	(0.06)	(2,330.77)	(2,412.30)	(0.02)
Top	0.02	0.00	339.66	-56.07	-0.01
	(0.06)	(0.05)	(2,256.99)	(2,341.75)	(0.02)
Sample Mean	4.19	3.01	114,072.59	123,147.58	-0.03
N	22,948	22,948	22,948	22,948	10,570
Parish Type FE	Yes	Yes	Yes	Yes	Yes
Municipality FE	No	No	No	No	No

Notes: Standard errors in parentheses clustered at parish $\times$ immigration year level. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$. The estimates show how refugees' labor market outcomes from year of arrival to 2017 (Panel (a)), and 8-15 year upon immigration (Panel (b)), is affected by placement neighborhood type using linear regression. The dependent variables are: (1) cumulative years with any employment, (2) cumulative years of employment (full time equivalents), (3) cumulated labor income in USD (deflated to 2000-level), (4) cumulated business income in USD (deflated to 2000-level), (5) average task complexity if employed. Task complexity is the average value of cognitive and communicative task intensities relative to manual task intensity based on occupations merged to the O*NET skill index. The sample mean denotes the mean of the outcome considered in the top panel from year of immigration until 2018. We measure parish income groups one year prior to arrival based on median disposable in each parish among all parishes in Denmark in a given year In all regressions we control for individual characteristics observed at time of assignment by including controls for gender, marital status, family size, and country of origin as well as age and year fixed effects.

Table 10: Education Outcomes

	All Education	Basic	Vocational	Academic	Health Education
			(a) Ever		
Middle	0.016**	-0.000	0.017***	0.000	-0.001
	(0.008)	(0.002)	(0.006)	(0.005)	(0.004)
Top	0.022***	0.001	0.022***	0.001	0.003
	(0.008)	(0.002)	(0.006)	(0.006)	(0.004)
			(b) Within 8 years after immigration		
Middle	0.016**	-0.000	0.017***	-0.000	-0.000
	(0.008)	(0.002)	(0.006)	(0.005)	(0.004)
Top	0.023***	0.000	0.022***	0.001	0.004
	(0.008)	(0.002)	(0.005)	(0.006)	(0.004)
Sample Mean	0.15	0.01	0.09	0.07	0.05
N	22,948	22,948	22,948	22,948	22,948
Parish Type FE	Yes	Yes	Yes	Yes	Yes
Municipality FE	No	No	No	No	No

Notes: Standard errors in parentheses clustered at parish $\times$ immigration year level. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$. The regressions test if the probability of completing any of the education types after immigration dependent on initial neighborhood income group. The dependent variables are dummies indicating whether the refugee completed the formal education of the type considered from year of arrival until 2017 (Panel (a)), and within the first 8 years upon arrival (Panel (b)). We measure parish income groups one year prior to arrival based on median disposable in each parish among all parishes in Denmark in a given year. In all regressions we control for individual characteristics observed at time of assignment by including controls for gender, marital status, family size, and country of origin as well as age and year fixed effects. The sample mean denotes the mean of the outcome considered in the top panel from year of immigration until 2018.

Table 11: Mechanisms, Lifestyle Related Diseases

	Baseline	(1)	(2)	(3)	(4)	(5)	(6)
				(a) Ever diagnosed			
Middle	-0.019*	-0.019**	-0.023**	-0.019*	-0.018*	-0.020**	-0.018*
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Top	-0.019**	-0.019**	-0.015	-0.019**	-0.016*	-0.021**	-0.018*
	(0.009)	(0.009)	(0.011)	(0.010)	(0.010)	(0.010)	(0.011)
				(b) Diagnosed 8-15 years after immigration			
Middle	-0.016**	-0.018**	-0.016*	-0.016**	-0.016**	-0.017**	-0.014*
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Top	-0.018**	-0.020***	-0.013	-0.018**	-0.016**	-0.020***	-0.015*
	(0.007)	(0.007)	(0.009)	(0.007)	(0.007)	(0.008)	(0.008)
N	22,948	22,948	22,941	22,945	22,948	22,948	22,948
Parish Type FE	Yes	No	Yes	Yes	Yes	Yes	Yes
Municipality FE	No	No	Yes	No	No	No	No
Control	No	No	No	GP/Capita	Health	Co-Nationals	Poverty

Notes: Standard errors in parentheses clustered at parish $\times$ immigration year level. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$. The table tests potential mechanisms behind the estimated neighborhood effects by estimating model (2) with different sets of controls. In column (Baseline) we replicate the estimates from Table 4. In column (1) we exclude the parish fixed effects. In (2) we include municipality fixed effects, in (3) we include the number of GPs per capita in the municipality of assignment as a control, in (4) we include the logarithm of the number of incidences (share of inhabitants above 18) of lifestyle related diseases in the assignment municipality as a control. In column (5) we include the number and squared number of co-nationals in the neighborhood, and in column (6) we include the poverty rate in the neighborhood as a control. All municipality and neighborhood characteristics are measured one year prior to immigration. The coefficients on the controls in (3)-(6) are positive or virtually zero and insignificant. Only the controls in (4) are significant at the 5 percent level (with an estimated coefficient of 0.038 in Panel (a), and 0.029 in Panel (b)). In Panel (a) the dependent variable is an indicator for being diagnosed with a lifestyle related disease at some point from year of arrival until 2017. In Panel (b) the dependent variable is an indicator for being diagnosed with a lifestyle related disease 8-15 years after immigration. We measure parish income groups one year prior to arrival based on median disposable in each parish among all parishes in Denmark in a given year. All other parish characteristics are also measured one year prior to arrival. In all regressions we control for individual characteristics observed at time of assignment by including controls for gender, marital status, family size, and country of origin as well as age and year fixed effects.

Table 12: IV Estimates, Different Definitions of Neighborhoods

	Lifestyle Related	Circulatory	Nutritional	Hypertension	Diabetes	Mental Disorder	Neurotic
				(a) Bottom Municipality			
Years of Exposure	0.004*	0.000	0.006***	0.002	0.003*	-0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
				(b) Bottom Parish			
Years of Exposure	0.003*	0.002	0.003**	0.002**	0.002*	0.000	0.000
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
				(c) Bottom Stairway			
Years of Exposure	0.012**	0.008	0.009**	0.006	0.009**	0.005	-0.000
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
N	18,581	18,581	18,581	18,581	18,581	18,581	18,581
Parish Type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	No	No	No	No	No	No	No

Notes: Standard errors in parentheses clustered at parish $\times$ immigration year level. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$. The table shows the increased probability of being diagnosed with one of the diseases in the top panel following an additional year living in a bottom neighborhood using different neighborhood definitions. In Panel (a) we let a municipality define a neighborhood and measure the increased probability of being diagnosed with the disease considered following an additional year spent in a bottom income municipality. We use initial municipality income group as instrument for the years spent in a bottom income municipality in the first stage. Completely parallel to that we let a parish define a neighborhood in Panel (b) using initial parish income group as an instrument in the first stage. Similarly, in Panel (c) we let a stairway (households living in the same building) define a neighborhood and use initial stairway income group as instrument in the first stage. In all neighborhood definitions we define neighborhood income groups based on median disposable income among adults of age 18 and above one year prior to arrival. The bottom income municipality, parish and stairway group refer to the bottom third of all municipalities, parishes and stairways, respectively. We measure income groups one year prior to arrival based on median disposable income in each year. In all regressions we control for individual characteristics observed at time of assignment by including controls for gender, marital status, family size, and country of origin as well as age and year fixed effects. F-statistics from first stage regression for years of exposure to bottom income municipality = 117.79, bottom income parish = 256.70, and bottom income stairway =78.57.

Table 13: Differences in Health Across Neighborhoods, Native Population

	Lifestyle Related	Circulatory	Nutritional	Hypertension	Diabetes	Mental Disorder	Neurotic
				(a) Ever diagnosed			
Middle	-0.009***	-0.005**	-0.008***	-0.004***	-0.003**	-0.009***	-0.005***
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Top	-0.016***	-0.007***	-0.014***	-0.005***	-0.006***	-0.024***	-0.012***
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.003)	(0.002)
Log Gross Income at	-0.010***	-0.004***	-0.008***	-0.001	-0.007***	-0.053***	-0.024***
Age 30	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
				(b) Diagnosed 8-15 years after 30th birthday			
Middle	-0.007***	-0.003**	-0.006***	-0.002**	-0.002***	-0.003***	-0.001*
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Top	-0.011***	-0.005***	-0.009***	-0.003***	-0.003***	-0.008***	-0.001**
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log Gross Income at	-0.008***	-0.003***	-0.005***	-0.001***	-0.003***	-0.016***	-0.002***
Age 30	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Sample Mean	0.23	0.15	0.10	0.07	0.04	0.16	0.09
N	384,250	384,250	384,250	384,250	384,250	384,250	384,250
Parish Type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors in parentheses clustered at parish × year level. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$. Estimates from a linear probability model testing the impact of the parish in which the individual lived at age 30 on the probability of being diagnosed with each of the diseases in the top panel. The sample consists of all individuals, disregarding the refugees, living in Denmark between 1994 and 1998 who turned 30 in that period. Income groups refer to income group of the parish in which the individual lived at age 30. The estimates show the the increased probability of being diagnosed with each of the considered diseases if living in the middle or top third income neighborhoods compared to the bottom third income neighborhoods. In Panel (a) the dependent variable is a dummy for being diagnosed with the disease considered at some point from year of 30$^{\text{th}}$ birthday until 2017. In Panel (b) the dependent variable is a dummy for being diagnosed with the disease considered 8-15 years after year of 30$^{\text{th}}$ birthday. We measure parish income groups one year prior to arrival based on median disposable income in each parish among all parishes in Denmark in a given year. Besides the fixed effects we control for marital status, gender and year fixed effects for year in which the individual turned 30. The sample mean denotes the share of natives diagnosed with the disease in the top panel before 2018.

References

Acemoglu, Daron and Joshua Angrist. 2000. "How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws." *NBER Macroeconomics Annual* 15:9–59.

Angrist, Joshua and Alan Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *The Quarterly Journal of Economics* 106 (4):979–1014.

Angrist, Joshua and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Bago d'Uva, Teresa and Andrew M. Jones. 2009. "Health care utilisation in Europe: New evidence from the ECHP." *Journal of Health Economics* 28 (2):265–279.

Boje-Kovacs, Bence, Jane Greve, and Cecilie Dohlmann Weatherall. 2018. "Can a shift of neighborhoods affect mental health? Evidence from a quasi-random allocation of applicants in the public social housing system." *Kraks Fond Institute for Urban Economic Research.* MPRA Paper No. 88929.

Chetty, Raj, Nathaniel Hendren, and Lawrence F Katz. 2016. "The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment." *American Economic Review* 106 (4):855–902.

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler. 2016. "The association between income and life expectancy in the United States, 2001-2014." *JAMA* 315 (16):1750–1766.

Christakis, Nicholas A. and James H. Fowler. 2007. "The Spread of Obesity in a Large Social Network over 32 Years." *New England Journal of Medicine* 357 (4):370–379. PMID: 17652652.

Damm, Anna Piil. 2004. "The Danish Dispersal Policy on Refugee Immigrants 1986-1998: A Natural Experiment?" *Working Paper 05-3* .

———. 2009. "Ethnic Enclaves and Immigrant Labor Market Outcomes: Quasi-Experimental Evidence." *Journal of Labor Economics* 27 (2):281–314.

———. 2014. "Neighborhood Quality and Labor Market Outcomes: Evidence from Quasi-Random Neighborhood Assignment of Immigrants." *Journal of Urban Economics* 79:139–166.

Damm, Anna Piil and Christian Dustmann. 2014. "Does Growing Up in a High Crime Neighborhood Affect Youth Criminal Behavior?" *American Economic Review* 104 (6):1806–32.

Danish Refugee Council. 1991. "Flygtninge i almennyttigt boligbyggeri." :1–10. Boligselskabernes Landsforening.

———. 1996. "Dansk Flygtningehjælps Integrationsarbejde." :1–64. Den Centrale Integrationsafdeling, September 12, 1996.

Dustmann, Christian, Kristine Vasiljeva, and Anna Piil Damm. 2018. "Refugee Migration and Electoral Outcomes." *The Review of Economic Studies* 86 (5):2035–2091.

Eisenberg, Daniel, Ezra Golberstein, Janis Whitlock, and Marilyn F Downs. 2013. "Social Contagion Of Mental Health: Evidence From College Roommates." *Health Economics* 22.

Fadlon, Itzik and Torben Heien Nielsen. 2019. "Family Health Behaviors." *American Economic Review* 109 (9):3162–91.

Finkelstein, Amy, Matthew Gentzkow, and Heidi L Williams. 2019. "Place-Based Drivers of Mortality: Evidence from Migration." *NBER Working Paper No. 25975* .

Foged, Mette and Giovanni Peri. 2015. "Immigrants' Effect on Native Workers: New Analysis on Longitudinal Data." *American Economic Journal: Applied Economics* 8 (2):1–34.

Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman. 2001. "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment*." *The Quarterly Journal of Economics* 116 (2):607–654.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1):83–119.

Ludwig, Jens, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, and Lisa Sanbonmatsu. 2012. "Neighborhood Effects on the Long-Term Well-Being of Low-Income Adults." *Science* 337 (6101):1505–1510.

Ludwig, Jens, Lisa Sanbonmatsu, Lisa Gennetian, Emma Adam, Greg J. Duncan, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, Stacy Tessler Lindau, Robert C. Whitaker, and Thomas W. McDade. 2011. "Neighborhoods, Obesity, and Diabetes A Randomized Social Experiment." *New England Journal of Medicine* 365 (16):1509–1519. PMID: 22010917.

Ministry for Cities, Housing and Rural Areas. 2013. "Regional- og Landdistriktspolitisk Redegoerelse 2013."

Oreopoulos, Philip. 2003. "The Long-Run Consequences of Living in a Poor Neighborhood." *The Quarterly Journal of Economics* 118 (4):1533–1575.

Patienthåndbogen. 2017. "Livsstilssygdomme." URL https://www.sundhed.dk/borger/patienthaandbogen/hjerte-og-blodkar/sygdomme/diverse/livstilssygdomme/.

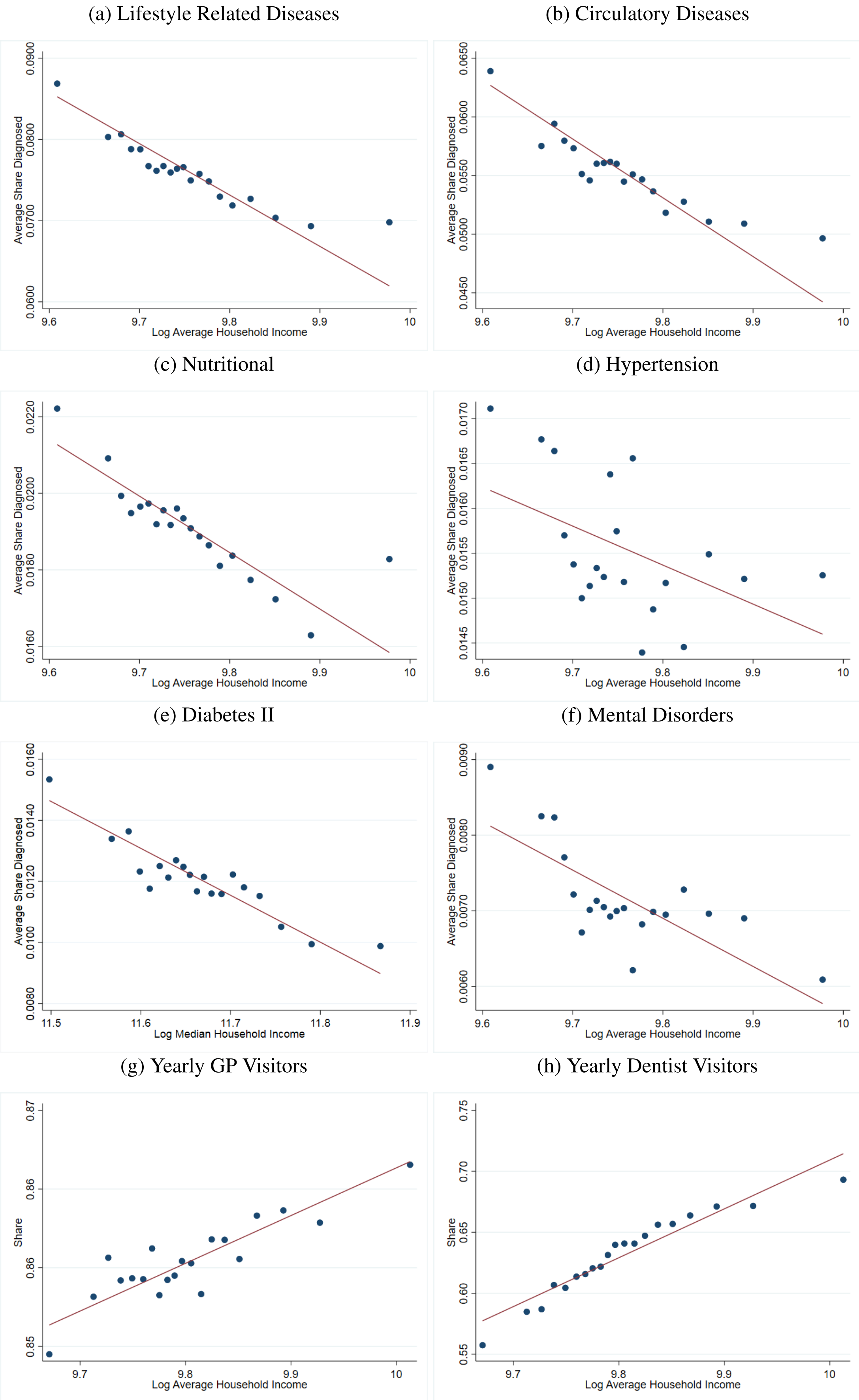
Sanbonmatsu, Lisa, Jens Ludwig, Lawrence F. Katz, Lisa A. Gennetian, Greg J. Duncan, Ronald C. Kessler, Emma Adam, Thomas W. McDade, and Stacy Tessler Lindau. 2011. "Moving to Opportunity for Fair Housing Demonstration Program – Final Impacts Evaluation."

White, Justin, Rita Hamad, Xinjun Li, Sanjay Basu, Henrik Ohlsson, Jan Sundquist, and Kristina Sundquist. 2016. "Long-Term Effects of Neighbourhood Deprivation on Diabetes Risk: Quasi-Experimental Evidence from a Refugee Dispersal Policy in Sweden." *The Lancet Diabetes & Endocrinology* 4.

World Health Organization. 2018. "Noncommunicable diseases." URL https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases.

A Appendix: Additional Tables and Figures

Figure A.1: Association Between Health and Neighborhood Income



Notes: The figures illustrate the association between health, health behavior and income between parishes. Panels a-f plot the average share in a parish diagnosed with the disease in question against the parish median disposable income, averaged over 1991-2017. Panels g-h plot the average share of inhabitants in a parish that visited their GP or dentist, respectively, against the parish median disposable income, averaged over 1991-2017. These unconditional correlations do not account for any selection or differences in inhabitant composition such as age or gender across parishes. Data stems from administrative data provided by Statistics Denmark from 1991-2017 for the full Danish population above 18 year of age.

Table A.1: Balancing Tests, Stairway Level

	(1) Bottom Income Group	(2) Middle Income Group	(3) Top Income Group	(4) Lifestyle Related
Unobserved at Time of Allocation				
Unknown Education	0.006	0.002	-0.009	0.000
	(0.015)	(0.015)	(0.011)	(0.002)
Basic Education	0.012	0.007	-0.019	0.003
	(0.016)	(0.016)	(0.012)	(0.002)
Academic Education	0.017	-0.012	-0.005	0.003
	(0.019)	(0.018)	(0.014)	(0.003)
Circulatory Disease	-0.008	-0.003	0.011	-0.002
	(0.032)	(0.033)	(0.024)	(0.004)
Nutritional Disease	-0.043	-0.014	0.056	-0.001
	(0.042)	(0.042)	(0.035)	(0.005)
Neurotic Disorder	-0.010	-0.074	0.085	-0.001
	(0.090)	(0.079)	(0.073)	(0.015)
Observed at Time of Allocation				
Age 30-49 Years	0.003	-0.013	0.011	-0.002
	(0.009)	(0.008)	(0.006)	(0.001)
Age 50-64 Years	-0.070***	0.063***	0.007	-0.001
	(0.014)	(0.015)	(0.011)	(0.002)
Female	-0.061***	0.041***	0.020***	0.003***
	(0.006)	(0.006)	(0.004)	(0.001)
Number of Adults	-0.031***	0.003	0.029***	0.003
	(0.011)	(0.011)	(0.009)	(0.002)
Number of Children	-0.022	0.014	0.008	-0.001
0-2 Years Old	(0.013)	(0.013)	(0.010)	(0.002)
Number of Children	-0.014***	0.003	0.011***	-0.000
3-17 Years Old	(0.004)	(0.004)	(0.003)	(0.001)
Married	-0.040***	0.037***	0.003	0.001
	(0.009)	(0.009)	(0.006)	(0.001)
Year of Immigration FE	Yes	Yes	Yes	Yes
Country of Origin FE	Yes	Yes	Yes	Yes
Parish Type FE	No	No	No	No
Municipality FE	No	No	No	No
N	20,804	20,804	20,804	20,806
F	0.37	0.40	1.28	0.83
Pr > F	0.90	0.88	0.26	0.55

Notes: Balancing tests for stairways using linear regressions. Standard errors in parentheses clustered at the household level. $^{**}p < 0.05,^{***} p < 0.01$. F denotes the F-statistic for joint insignificance of the educational attainment dummies and pre-existing health conditions. Each column represents a different balancing test testing if refugees with certain characteristics (column farthest to the left) are more likely to be placed in stairways with specific characteristics (dependent variables). The dependent variables in (1)-(3) are dummies for assignment to a bottom income stairway (1), middle income stairway (2) or top income stairway (3). In column (4) the dependent variable is the incidence (as a share of inhabitants) of lifestyle related diseases. The controls are individual characteristics observed by the DRC at time of assignment and characteristics which the DRC does not observe at time of assignment: initial education and initial health. As a proxy for initial health we use diagnoses within the first year upon arrival, but measure all other individual characteristics at year of immigration. We measure all stairway characteristics one year prior to immigration.

Table A.2: Balancing Tests, Municipality Level

	(1) Bottom Income Group	(2) Middle Income Group	(3) Top Income Group	(4) Lifestyle Related
Unobserved at Time of Allocation				
Unknown Education	0.004	0.017	-0.021	-0.000
	(0.009)	(0.014)	(0.013)	(0.000)
Basic Education	0.007	0.028	-0.036**	-0.000
	(0.010)	(0.015)	(0.014)	(0.000)
Academic Education	0.006	0.018	-0.024	0.000
	(0.012)	(0.017)	(0.017)	(0.000)
Circulatory Disease	-0.002	0.020	-0.018	0.000
	(0.020)	(0.030)	(0.029)	(0.000)
Nutritional Disease	-0.013	0.058	-0.045	0.000
	(0.027)	(0.039)	(0.036)	(0.000)
Neurotic Disorder	-0.086	0.055	0.032	-0.001
	(0.050)	(0.074)	(0.075)	(0.001)
Observed at Time of Allocation				
Age 30-49 Years	-0.013**	0.011	0.002	-0.000
	(0.005)	(0.008)	(0.007)	(0.000)
Age 50-64 Years	-0.024***	0.032**	-0.008	-0.000
	(0.008)	(0.013)	(0.013)	(0.000)
Female	-0.009**	-0.012**	0.021***	-0.000
	(0.004)	(0.006)	(0.005)	(0.000)
Number of Adults	-0.002	-0.012	0.014	-0.000***
	(0.006)	(0.010)	(0.010)	(0.000)
Number of Children	0.013	-0.005	-0.008	-0.000
0-2 Years Old	(0.008)	(0.012)	(0.011)	(0.000)
Number of Children	-0.002	0.003	-0.001	-0.000***
3-17 Years Old	(0.002)	(0.003)	(0.003)	(0.000)
Married	0.013**	-0.010	-0.003	-0.000
	(0.005)	(0.008)	(0.008)	(0.000)
Year of Immigration FE	Yes	Yes	Yes	Yes
Country of Origin FE	Yes	Yes	Yes	Yes
Parish Type FE	No	No	No	No
Municipality FE	No	No	No	No
N	25,738	25,738	25,738	25,738
F	0.64	1.22	1.52	1.00
Pr > F	0.70	0.29	0.17	0.42

Notes: Balancing tests for municipalities using linear regressions. Standard errors in parentheses clustered at the household level. $^{**}p < 0.05, ^{***}p < 0.01$. F denotes the F-statistic for joint insignificance of the educational attainment dummies and pre-existing health conditions. Each column represents a different balancing test testing if refugees with certain characteristics (column farthest to the left) are more likely to be placed in municipalities with specific characteristics (dependent variables). The dependent variables in (1)-(3) are dummies for assignment to a bottom income municipality (1), middle income municipality (2) or top income municipality (3). In column (4) the dependent variable is the incidence (as a share of inhabitants) of lifestyle related diseases. The controls are individual characteristics observed by the DRC at time of assignment and characteristics which the DRC does not observe at time of assignment: initial education and initial health. As a proxy for initial health we use diagnoses within the first year upon arrival, but measure all other individual characteristics at year of immigration. We measure all municipality characteristics one year prior to immigration.

A.1 Diagnoses with ICD Codes

The first parentheses indicate (ICD-10) diagnoses codes from 1994 and onwards and second parentheses indicate (ICD-8) diagnoses codes before 1994. Diagnoses with bold correspond to the groups we use in our regression analysis.

Lifestyle related diseases:

- **Circulatory diseases:**
  - **Hypertensive diseases** (referred to as hypertension): (I10), (400-401)
  - Ischaemic heart diseases: (I20, I22, I24, I25), (411-414)
  - Pulmonary diseases: (I26-I28), (426, 450, 514)
  - Other forms of heart diseases: (I30-I52), (393-398, 420-429)
  - Cerebrovascular diseases: (I60-I67, I69), (430-438)
  - Arterial diseases: (I70-I72,I74), (440-442, 444)
- **Endocrine, nutritional and metabolic diseases** (referred to as nutritional diseases):
  - **Diabetes**: (E10-E14), (250)
  - Obesity: (E66), (277)
  - Metabolic disorders (high cholesterol): (E78), (272)
- Chronic obstructive pulmonary diseases (COPD): (J44), (490, 491, 492)
- Hip arthrosis: (M16), (710.2)
- Alcohol related diseases:
  - Alcohol induced acute pancreatitis: (K85.2), (577.0),
  - Alcoholic liver disease: (K70), (571.0)
  - Alcoholism: (No ICD10 code), (303)

Mental disorders:

- Mental and behavioral disorders due to psychoactive substance use: (F10-F19), (291, 294.3, 309.1, 29430, 29438, 29439, 30919)
- Schizophrenia, schizotypal and delusional disorders: (F20-F29), (295)
- Mood [affective] disorders: (F30-F39), (296)
- **Neurotic, stress-related and somatoform disorders:** (F40-F48), (300)
- Behavioral syndromes associated with physiological disturbances and physical factors: (F50-F59), (305)
- Disorders of adult personality and behavior: (F60-F69), (301, 302)

Congenital disorders:

- **Congenital abnormalities:** (Q00-Q99), (740-759)
- **Congenital metabolic disorders:** (E70-E77, E79-E90), (270-271, 273-276, 278-279)

Table A.3: Impact on Mortality

	Within 8 YSM	Within 15 YSM	Before 2018
Middle	-0.002	-0.007	-0.007
	(0.005)	(0.006)	(0.007)
Top	-0.006	-0.009	-0.007
	(0.005)	(0.006)	(0.006)
Sample Mean	0.04	0.07	0.10
N	22,948	22,948	22,948
Parish Type FE	Yes	Yes	Yes
Municipality FE	No	No	No

Note: Standard errors in parentheses clustered at parish $\times$ immigration year level. $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$. The estimates show the increased probability of death if assigned to a middle- or a top-income neighborhood compared to a bottom income neighborhood. In the first column the dependent variable is a dummy for dying within the first 8 years after immigration, in the second column the dependent variable is dummy for dying within the first 15 years since immigration. In the last column the dependent variable is a dummy for dying before 2018. We measure parish income groups one year prior to arrival based on median disposable income in each parish among all parishes in Denmark in a given year. In all regressions we control for individual characteristics observed at time of assignment by including controls for gender, marital status, family size, and country of origin as well as age and year fixed effects. The sample mean denotes the mean of the outcome considered in the top panel from year of immigration until 2018.

Table A.4: Mechanisms, Lifestyle Related Diseases

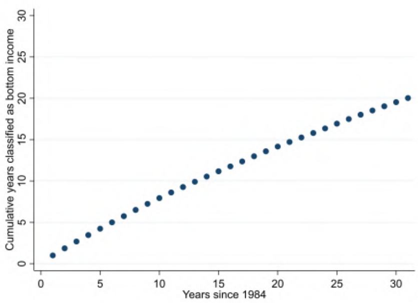
	Baseline	(1)	(2)	(3)	(4)
			(a) Ever diagnosed		
Middle	-0.019*	-0.019*	-0.018*	-0.019*	-0.019*
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Top	-0.019**	-0.018*	-0.020**	-0.020**	-0.020**
	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)
			(b) Diagnosed 8-15 years after immigration		
Middle	-0.016**	-0.016**	-0.016**	-0.017**	-0.018**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Top	-0.018**	-0.017**	-0.018**	-0.018**	-0.019**
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
N	22,948	22,945	22,948	22,948	22,866
Parish Type FE	Yes	Yes	Yes	Yes	Yes
Municipality FE	No	No	No	No	No
Control	No	Health Expenditure	Number Refugees	Immigrant Share	Immigrant Income

Notes: Standard errors in parentheses clustered at parish $\times$ immigration year level. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$. The table tests potential mechanisms driving the estimated neighborhood effects by estimating the increased probability of being diagnosed with a lifestyle related disease following assignment to a middle- or top-income neighborhood compared to bottom-income neighborhoods with different sets of controls. In column (Baseline) we repeat the estimates from Table 4. In column (1) we include the control "Health Expenditure", which refers to the inclusion of the logarithmic number of GPs per capita in the municipality and the logarithmic health and social expenditures per capita in the municipality. In column (2) we control for the number of refugees by including the number of inhabitants in the neighborhood originating from any of the refugee sending countries in our sample as a control. In column (3) we include the share of immigrants and the squared share of immigrants as a control. In column (4) we include the logarithmic of median disposable income among immigrants in the neighborhood. We measure all these neighborhood controls one year prior to arrival. In Panel (a) the dependent variable is a dummy for being diagnosed with a lifestyle related disease before 2018, and in Panel (b) the dependent variable is a dummy for being diagnosed with a lifestyle related disease between 8-15 years after immigration. We measure parish income groups one year prior to arrival based on median disposable income in each parish among all parishes in Denmark in a given year. All other parish characteristics are also measured one year prior to arrival. In all regressions we control for individual characteristics observed at time of assignment by including controls for gender, marital status, family size, and country of origin as well as age and year fixed effects.

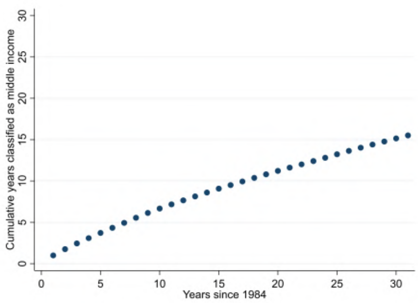
A.2 Instrumental variables strategy

Figure A.2: Persistence in Neighborhood Classifications

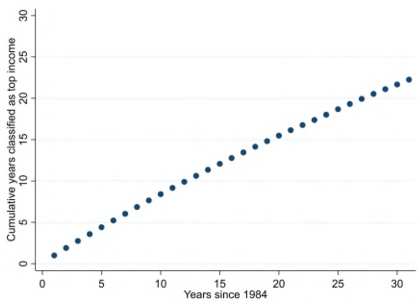
(a) Bottom Income Parish



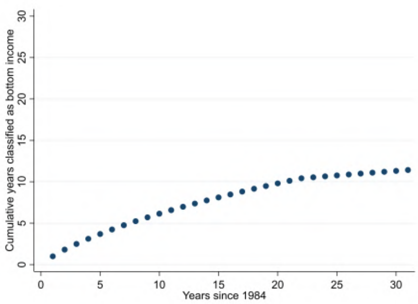
(b) Middle Income Parish



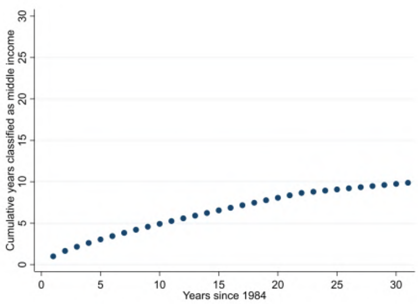
(c) Top Income Parish



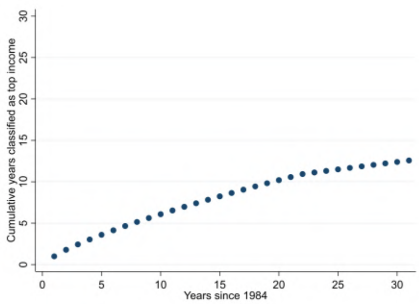
(d) Bottom Income Stairway



(e) Middle Income Stairway



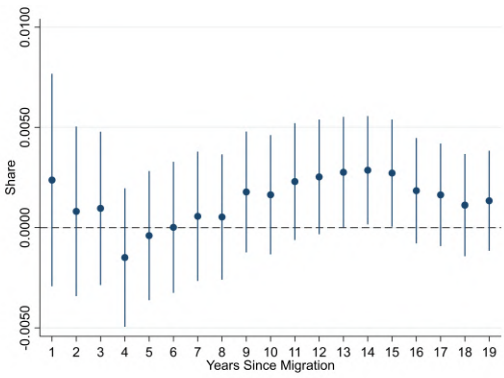
(f) Top Income Stairway



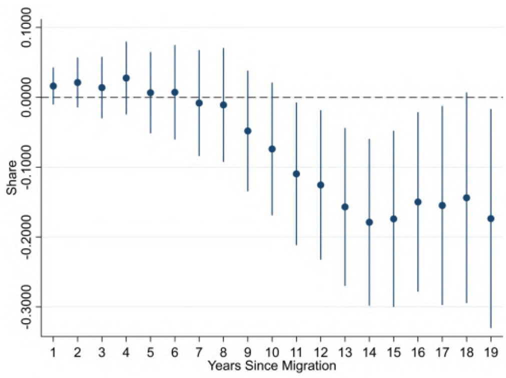
Notes: Panel (a) shows the cumulative number of years a parish, which belongs to the bottom third income group of parishes in 1984, belongs to the bottom third income group of parishes until 2017 measured by median disposable income among adults in each parish in each year. Similarly, Panels (b) and (c) show this for parishes originally classified as the middle or top third income groups of parishes, respectively. Panels (d), (e) and (f) show the exact same for stairways.

Figure A.3: Dynamics of Lifestyle Related Diagnoses, IV Estimates

(a) Exposure to Bottom Income Neighborhoods



(b) Impact of Average Neighborhood Income



Notes: In Panel (a) cumulative years of exposure to bottom income neighborhoods is instrumented by placement neighborhood income group. In Panel (b) the average income in all neighborhoods lived in until year $t + r$ is instrumented by the average income in the first placement neighborhood. Standard errors clustered at parish $\times$ immigration year level. 90 percent confidence intervals. In all regressions we control for individual characteristics observed at time of assignment by including controls for gender, marital status, family size, and country of origin as well as age and year fixed effects.

A.3 Alternative instrumental variables strategy

Another approach to take endogenous moving into account, is to instrument the average income level that the refugee was exposed to over the r years since arrival. As an instrument we use the initial income level in the placement neighborhood one year prior to arrival. We then estimate the effect of experiencing a higher average neighborhood income level since arrival. In this approach we calculate the average income level of all neighborhoods which the refugee lived in during the r years after arrival: $\bar{x}_{i,t+r} \equiv \frac{\sum_{r=0}^{r} income_{n,t+r}}{r}$. We instrument this average with the income level of the placement neighborhood at time $t-1$. Again, this instrument is relevant if there is some persistence in neighborhood income levels experienced after arrival. If this is fulfilled we can estimate:

$$Second\ stage: \quad y_{i,t+r} = \alpha_1 + \beta_1 \hat{\bar{x}}_{i,t+r} + \boldsymbol{X_{it}\gamma_1} + \boldsymbol{T_t} + \boldsymbol{P_n} + \varepsilon_{i,t+r} \tag{5}$$

In model (5), $income_{n,t-1}$ denotes income in initial placement neighborhood one year prior to arrival and $\hat{\bar{x}}_{i,t+r}$ denotes the average neighborhood income level experienced over the r years since arrival. All other inputs are the same as in models (2) and (3). The coefficient β_1 can be interpreted as the increased risk of being diagnosed with y when living in neighborhoods with one percent higher income for r years.

Table A.5: IV Results, Average Neighborhood Income Level

	(1) Lifestyle Related	(2) Circulatory	(3) Nutritional	(4) Hypertension	(5) Diabetes	(6) Mental Disorder	(7) Neurotic
Average Income	-0.185*	-0.068	-0.085	-0.154**	-0.121	-0.132	-0.039
	(0.102)	(0.086)	(0.085)	(0.068)	(0.074)	(0.099)	(0.085)
N	22,948	22,948	22,948	22,948	22,948	22,948	22,948
Parish Type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	No	No	No	No	No	No	No

Notes: Standard errors in parentheses clustered at parish $\times$ immigration year level. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$. The tables shows the increased probability of being diagnosed with each of the diseases considered in the top panel following an increase in average income in neighborhoods lived in since immigration of 1 pct. The average neighborhood income level in all neighborhoods lived in since immigration is instrumented by the median neighborhood income among adults of age 18 and above in the first placement neighborhood. We control for individual characteristics observed at time of assignment by including controls for gender, marital status, family size, and country of origin as well as age and year fixed effects. F-statistics from first stage regression for average income in parishes lived in is =121.30

A.4 Exposure to the Poorest Neighborhoods, Without IV

Table A.6: Exposure to the Poorest Neighborhoods, OLS

	Lifestyle Related	Circulatory	Nutritional	Hypertension	Diabetes	Mental Disorder	Neurotic
Years of Exposure	0.007***	0.004***	0.006***	0.003***	0.004***	0.007***	0.005***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	23,141	23,141	23,141	23,141	23,141	23,141	23,141
Parish Type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	No	No	No	No	No	No	No
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS

Notes: Standard errors in parentheses clustered at parish $\times$ immigration year level. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$. The table shows the increased risk of being diagnosed with one of the diseases in the top panel following an additional year of exposure to a bottom income bottom neighborhood. We measure parish income groups one year prior to arrival based on median disposable in each parish among all parishes in Denmark in a given year. In all regressions we control for individual characteristics observed at time of assignment by including controls for gender, marital status, family size, and country of origin as well as age and year fixed effects.

A.5 Neighborhood definition

Table A.7: IV With Multiple Instruments

	Lifestyle Related	Circulatory	Nutritional	Hypertension	Diabetes	Mental Disorder	Neurotic
Years of Exposure to	-0.001	0.001	-0.001	0.001	-0.001	-0.000	0.002
Bottom Parish	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)
Years of Exposure to	0.010*	0.002	0.009	0.003	0.009*	0.002	-0.002
Bottom Stairway	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)	(0.006)	(0.005)
N	18,424	18,424	18,424	18,424	18,424	18,424	18,424
Parish Type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	No	No	No	No	No	No	No

Notes: Standard errors clustered at parish $\times$ immigration year level. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$. The table shows the increased probability of being diagnosed with one of the diseases in the top panel following an additional year spent in a bottom income neighborhood using different neighborhoods at the parish and stairway level simultaneously. We use initial parish and stairway income groups as instruments for years spent in a bottom income parish and stairway in the first stage. In both neighborhood definitions we define neighborhood income groups based on median disposable income among adults of age 18 and above one year prior to arrival. The bottom income parish and stairway group refer to the bottom third of all parishes and stairways, respectively. We measure income groups one year prior to arrival based on median disposable in each year. In all regressions we control for individual characteristics observed at time of assignment by including controls for gender, marital status, family size, and country of origin as well as age and year fixed effects. F-statistics from first stage regression for years of exposure to bottom parish and stairway is =133.19

Table A.8: Summary Statistics for Initial Placement (Stairway)

	Bottom Mean	Middle Mean	Top Mean
Characteristics of Residents			
Age	40.28	39.57	38.81
Median Household Income	13,563.90	14,221.27	14,855.93
Employment Rate	0.48	0.55	0.59
Prevalence of Lifestyle Related Diseases per 1,000 Inhabitants	69.90	67.45	47.69
Inhabitants	20.92	11.32	13.48
Co-Nationals	1.26	0.83	0.77
Poverty Rate	0.13	0.10	0.10
Parish Type			
Urban Area (Near City)	0.45	0.43	0.68
Urban Area (Away from City)	0.04	0.19	0.16
Rural Area (Near City)	0.09	0.10	0.08
Rural Area (Away from City)	0.30	0.21	0.05
Characteristics of Municipality			
General Practioners per 1,000 Inhabitants	0.46	0.43	0.46
Incidences of Lifestyle Related Diseases per 1,000 Inhabitants	33.01	29.31	26.11
Health and Social Expenditures per Capita	4,016.16	4,112.72	4,022.29
N	683	1,456	2,773

Notes: Summary statistics for stairways in which refugees were resettled. A stairway refers to the group of households living in the same building. ”Bottom”, ”Middle” and ”Top” refer to stairway characteristics of stairways in the bottom, middle and top third of stairways measured by median stairway disposable income in a given year. We calculate the median income of each stairway including all inhabitants in each stairway aged 18 or above and define the income groups among all stairways, irrespective of any refugee assignment. We define income groups and all stairway characteristics one year prior to immigration. Prevalence of lifestyle related diseases is measured as all incidences over the previous 8 years and thus only defined for refugees arriving after 1993. Employment rate is the share of the population with any employment between the ages of 18-65. Observations are stairway-year. Health and social expenditures per capita and median household income is measured in USD.

Table A.9: Summary Statistics for Initial Placement (Municipality)

	Bottom Mean	Middle Mean	Top Mean
Characteristics of Residents			
Age	47.91	47.46	45.98
Median Household Income	14,610.13	14,678.44	15,936.67
Employment Rate	0.67	0.69	0.73
Inhabitants	30,743.78	20,078.51	22,332.97
Co-nationals	54.05	37.06	27.10
Poverty Rate	0.08	0.07	0.06
Urban Area (Near City)	0.15	0.23	0.58
Urban Area (Away from City)	0.09	0.29	0.22
Rural Area (Near City)	0.18	0.14	0.11
Rural Area (Away from City)	0.54	0.31	0.06
Characteristics of Municipality			
General Practioners per 1,000 Inhabitants	0.37	0.36	0.41
Incidences of Lifestyle Related Diseases per 1,000 Inhabitants	32.40	29.24	24.65
Health and Social Expenditures per Capita	3,562.66	3,648.68	3,552.23
N	199	511	1,021

Notes: Summary statistics for municipalities in which refugees were resettled. ”Bottom”, ”Middle” and ”Top” refer to municipality characteristics of municipalities in the bottom, middle and top third of municipalities measured by median municipality disposable income in a given year. We calculate the median income of each municipality including all inhabitants in each municipality aged 18 or above and define the income groups among all municipalities, irrespective of any refugee assignment. We define income groups and all municipality characteristics one year prior to immigration. Employment rate is the share of the population with any employment between the ages of 18-65. Observations are municipality-year. Health and social expenditures per capita and median household income is measured in USD.

Chapter 2

Integrating Refugees: The Role of Language Training and Work Incentives

INTEGRATING REFUGEES:

THE ROLE OF LANGUAGE TRAINING AND WORK INCENTIVES*

Jacob Nielsen Arendt[†], Iben Bolvig[‡], Mette Foged[§], Linea Hasager[¶] and Giovanni Peri[||]

August 12, 2020

Abstract

Social and economic integration of refugees are key to their personal success and to producing positive effects in the host country. We evaluate the effects of a reform that substantially expanded language training for immigrants who obtained refugee status in Denmark on or after January 1, 1999. The same reform also temporarily decreased welfare benefits for a subgroup of them. Using a regression discontinuity design around the cutoff date we find positive and significant employment and earnings effects on the treated group, relative to the untreated group. Employment increased by 23 percent (4 percentage points) and yearly earnings increased by 34 percent (USD 2,500) when measured eighteen years after the start of the language program. We do not find any labor market effect of the reduction of welfare benefits. We find, however, evidence of temporarily higher property crimes for the group subject to lower benefits. The labor market effects are much stronger for individuals with Arabic/Dari mother language, consistently with a more crucial role of language training for speakers of languages that are very different from Danish. Finally, male children of treated refugees were more likely to complete lower secondary school and less likely to commit crime.

JEL Classification: J60, J24, E64, I30.
Keywords: Refugee Integration, Language Skills, Welfare Benefits, Regression Discontinuity, Second Generation.

*Corresponding author: Mette Foged. This project is conducted as part of the Economic Assimilation Research Network (EARN), generously financed by the Innovation Fund Denmark (grant no. 6149-00024B). Mette Foged also acknowledges funding from EPRN (grant no. 107757). We thank the Ministry of Immigration and Integration for assistance with data and reports on early language training in Denmark and Janis Kreuder for excellent research assistance. The project benefited from feedback and discussions with Oddbjørn Raaum, Jakob Roland Munch, Torben Tranæs, Matti Sarvimäki and participants at the following seminars and workshops: "Forced Displacement, Asylum Seekers and Refugees", Queen Mary University London; the first "EARN Workshop on Integration", University of Copenhagen; Uppsala; Global Migration Research Center, UC Davis; Danish Center for Social Science Research (VIVE); CReAM, UCL; University of Bristol; Senior Migration Seminar (`sites.google.com/view/the-economics-of-migration/senior-seminar`); and FAIR, NHH.

[†]The Rockwool Foundation's Research Unit, Ny Kongensgade 6, DK 1472 Copenhagen K, jar@rff.dk.
[‡]The Danish Center for Social Science Research (VIVE), Olof Palmes Alle 22, DK 8200 Aarhus N, IbBo@vive.dk.
[§]University of Copenhagen, Øster Farimagsgade 5 Building 26, DK 1353 Copenhagen, Mette.Foged@econ.ku.dk.
[¶]University of Copenhagen, Øster Farimagsgade 5 Building 26, DK 1353 Copenhagen, tlh@econ.ku.dk.
[||]University of California, Davis, One Shields Avenue, Davis CA 95616, gperi@ucdavis.edu.

I Introduction

International refugees, who usually escape from extreme conditions of persecution or war, begin their lives in the host country with significant disadvantages. Research shows that even in the medium and long run, the employment and wage gap between them and natives or other immigrants is significant (Fasani, Frattini, and Minale, 2018; Brell, Dustmann, and Preston, 2020).

Policies that are effective at improving the labor market integration of refugees, therefore, could bring large economic returns to the migrants and to their host economy. These economic benefits make other forms of social integration more likely, possibly reducing segregation and crime.

While the goal of enhancing the labor market outcomes of refugees is a high priority of host countries, the way to effectively achieve such outcomes is less clear. Integration efforts, therefore, have taken many forms. Several countries, such as the US, leave this effort to local interventions (Bloemraad and De Graauw, 2012; Williamson, 2018), often coordinated by private charities and churches. Most European governments, instead, consistently with their tendency to intervene more in the economy and society, have supported and funded early interventions for refugees since the 1990s (Arendt, 2018). Integration programs are usually a combination of several forms of intervention, including active labor market support, provision of incentives, and teaching of relevant skills. How effective each of these interventions are in improving labor market and social outcomes for refugees has been subject to a lively debate but to very limited rigorous analysis. Limited longitudinal data following refugees over time and lack of policies with a clean experimental design have prevented a causal assessment of those interventions.

Lacking a consensus on how effective each policy intervention is, different countries have moved in different directions. Germany and France have increased their focus on immigrants' language learning.[1] The Netherlands and other North-European countries, instead, moved in the opposite direction in recent years, reducing the focus on language skills while encouraging job search and labor market participation early on. While from 1998 to 2007 the Dutch government required new immigrants to participate in a 12-month language, social orientation and labour market orientation program (Joppke, 2007), starting in 2007 participation to this program was made optional.[2] Similarly, Denmark and other Northern European countries have recently moved away from policies focused on early language training and welfare assistance, which were prevalent in the 1990s and 2000s. In Denmark, since 2016 all refugees

[1] Germany mandated in 2005 language (and German history, law and culture) classes for newly arrived refugees (Hübschmann, 2015; Martin et al., 2016). France introduced the *Contract d'accueil et d'intgration* in 2007, requiring immigrants to participate in language training and civic training (Lochmann, Rapoport, and Speciale, 2019).

[2] Since 2013 the integration program has to be paid for by the immigrant herself (De Vries, 2013).

are required to actively search for jobs and participate in job-training immediately upon arrival, shifting the focus of integration interventions towards early labor market participation.[3]

As a consequence of these rather disparate integration efforts, interesting policy experiments have been implemented and, as administrative data availability has become more prevalent, we can use one of these policy experiments in Denmark to fill the gap in evaluating the causal role of enhanced language learning on economic integration. Specifically, the present paper sheds light on the causal effects of comprehensive language training on short and long-run outcomes of refugees and of their children. Additionally, as we will describe below, the policy considered allow us to learn about the impact of reducing welfare benefits.

Our identification strategy utilizes a reform of the Danish integration program to isolate the effect of access to more extensive language training in the early years after settlement. The reform involved a very significant increase in the duration and some additional improvements in the language training provided to refugees in Denmark. The integration program was mandatory for refugees and family members who were reunified to them (for brevity we refer to both as refugees). Individuals who obtained refugee status in Denmark on or after January 1, 1999 were entitled to the expanded language program. Those who obtained refugee status before that date did not qualify. Hence, we use a regression discontinuity (RD) design based on the date when refugee status was granted relative to January 1, 1999 which is the threshold date. The language training component was the more relevant and permanent part of the reform, and our primary focus is on evaluating its impact. Additionally, there was a smaller change which was also introduced at the same discontinuity date. This was a temporary but sizeable reduction of welfare benefits for refugees, limited to those over age 25 or with children. This second change was rolled back after 13 months (February 1, 2000). The welfare benefit reduction was conceived as a way of creating stronger job-finding incentives for newly arrived refugees in prime working age. As only a sub-group of refugees was subject to it, and as the policy was in place for only a year, we evaluate the potential effects of this additional treatment only given to individuals over 25 or with children, during the year 1999.

Several features make this reform interesting and the content of this paper novel. First, the introduction of this program with a sharp time discontinuity generates random assignment around the threshold date and allows us to credibly estimate a causal effect. Second, as we observe details about language training participation and subsequent education, we can analyze whether the treatment of extended lan-

[3] Arendt (2019) finds large positive employment effects of this new policy one year after settlement in Denmark. However he also finds that the effect is temporary and crowds out language course participation.

guage classes was the catalyst for further education and training. Third, as we can observe a rich set of variables for the refugees over an extended period of time, we can look at the effects on economic and social outcomes occurring both in the short and in the long run. Finally, as we can link the refugees to their spouses and children, we can analyze effects on family structure, fertility and on the second generation's outcomes. The long time horizon is unique to our paper in this literature. It is important because the success of an integration policy ultimately depends on the outcomes of refugees in the long run and, possibly, on the effects on the next generation.

Several important findings emerge from the analysis. First, being exposed to the expanded language training increases employment probabilities and earnings permanently.[4] The positive effects emerge four-five years after the reform was implemented and are statistically significant starting six years after the reform. Consistent with the fact that the language classes were attended in the first three years after settlement, we find zero effects on employment and earnings during the first three years, followed by a gradual improvement and then consistently higher earnings and employment in the treatment group up to eighteen years, the longest time horizon we can include.[5] Second, we find evidence of formal skill upgrading after the language classes, in the form of more years of schooling, consistently with the idea that better language skills facilitated access to education. We also see evidence of occupational upgrading to jobs with a stronger communication requirement.[6] Third, for the group subject to the additional temporary welfare cut (age 25 and older), we find a temporary increase in the probability of committing property crime (mainly shoplifting) exactly during the year when welfare was reduced. This effect on crime disappears in the following years when the original welfare payments were restored. Fourth, in a heterogeneity analysis we do not find significant differences by gender but we find differences across mother languages (reported through the language training facilities). The treatment had the largest employment effect on refugees who spoke Dari or Arabic (usually from Iraq and Afghanistan) and were not familiar with the Latin alphabet at arrival.[7] This suggests that refugees speaking languages very different from Danish may be benefiting more from the extensive language training. Finally, we do not find any impact of the reform on marriage outcomes or on fertility. However, when considering the children of refugees, we see that sons (but not daughters) of treated refugees were more likely to complete lower

[4]The average refugee increased language instruction time by 200 hours. The reform entitled the treated to 430 additional hours that could be taken over an extended period of time, three years compared to 1.5 years before the reform.

[5]The point estimates are extremely robust to alternative bandwidth and functional forms. The statistical significance of the earnings effects is also insensitive to most of the robustness checks we perform while the employment effects are on the margin of significance in some specifications.

[6]A simple cost-benefit analysis reveals a positive net present value of the additional language training with time to break-even never exceeding six years.

[7]We group all languages akin to the Arabic alphabet in the empirical analysis. The two languages mentioned here constitute 82 percent of individuals in this group.

secondary school and less likely to commit a crime when 16 years or older.

The main contributions and findings connect to two related strands of research. First, the literature on the role of language in the economic assimilation of immigrants. Early studies established a strong and positive association between language proficiency and the earnings of immigrants.[8] Part of this correlation certainly arises from omitted variable bias, as more able and more motivated immigrants acquire better language skills and earn more, and part of the correlation likely reflects a causal link between language proficiency and earnings in the host country (Borjas, 1994). Hence, a second and more recent strand of the literature tries to estimate the causal effect of language training and is focused on identifying the effect of language classes offered to refugees or other immigrants on their labor market outcomes and earnings. Two recent papers stand out in this literature. Closest to this study in its identification strategy is Sarvimäki and Hämäläinen (2016), that uses a reform in Finland affecting the integration services provided to immigrants who arrived after May 1, 1997, generating a discontinuity and random assignment near it. The authors show that the reform changed the composition of services provided during the first five years in Finland, towards more language training and away from traditional active labor market policies (such as job-seeking). The treated group had a very significant 47 percent increase in earnings and 13 percent decrease in cumulative social benefits over a ten year period but no change in employment. Our paper is different from Sarvimäki and Hämäläinen (2016) in several respects. First, the reform we consider is a clean change in language training (no change in regular active labor market support). Second, we focus on refugees, a particularly disadvantaged group with significant issues in economic integration. Third, we analyze outcomes besides employment and earnings, such as schooling and crime, which are quite relevant in designing policies for refugees' integration. Fourth, we analyze effects on the next generation. Finally, we analyze, as a secondary policy, the effect of cutting welfare temporarily on the sub-group of refugees over 25 years old or with children.

The second relevant paper focused on the impact of language training on economic outcomes of refugees is Lochmann, Rapoport, and Speciale (2019). They analyze the introduction of a program in France that assigns non-EU immigrants to language training based on a threshold in their score in an initial test.[9] Such a discontinuity in the score produces random assignment to the language treatment around the threshold. The authors find that 100 hours of language training increases labor force participation by 15 to 27 percentage points two years after completion of classes. This is a very large effect

[8] See for instance Chiswick (1991); Chiswick and Miller (1995) and Dustmann (1994).

[9] In France, differently from Denmark and Finland, many immigrants come from French-speaking countries, due to colonial ties. Furthermore, the reform was not restricted to the newly arrived and the language training provided was much less ambitious than programs in the Northern European countries.

equal to a quarter of the sample mean (labor force participation is 81 percent). However the authors also show some puzzling findings, such as no effect on employment (in spite of much larger participation) and no effect on the self-assessed ability to speak the French language. Moreover, they cannot analyze longer-term effects, nor many of the social and family outcomes we analyze, due to data limitations. The small number of observations that can be used in their RD design is a limitation of their paper and of the robustness of their results.

A few additional studies analyzing integration policies and immigrants' labor market outcomes are worth mentioning. LoPalo (2019) finds that the generosity of welfare benefits (Temporary Assistance for Needy Families) across US states had no significant effect on refugees' employment while it produced an increase in their wages in the long run. Identification is based on cross-state variation of benefit generosity and hence potentially subject to omitted variable bias. Existing studies relative to Denmark show that a substantial reduction in the welfare benefit for refugees, that occurred after reforms in 2002 and 2015, produced large positive short-term employment effects.[10] Andersen, Dustmann, and Landersø (2019) show that the effect disappeared after 7-8 years and that the same reform produced adverse effects on crime rates and educational achievements among teenagers in treated household.[11] Joona and Nekby (2012) evaluate a program in Sweden where new immigrants were randomly selected to receive more intensive counselling. They find significant employment effects one year after assignment but no evidence of longer-run effects. Taking the previous studies as a whole, the impact of active labor market policies and economic incentives on refugees labor market outcomes seems short-lived, while the returns to language training (documented by Sarvimäki and Hämäläinen and this paper) appear to be more long lasting.[12] The clean identification of additional language training, the long-run and second-generation evidence, detailed labor market and social integration outcomes and the ability to tackle the effects of language training and welfare transfer represent a significant new contribution of this paper relative to the studies mentioned above.

The rest of the paper is structured as follows. Section II discusses the recent history of refugees in Denmark. Section III describes the samples and the key outcomes we look at. Section IV describes the reform and its implementation. Section V discusses the empirical strategy. Section VI presents the main analysis of the working-age refugees, and section VII studies fertility and impacts on their children.

[10] See Huynh, Schultz-Nielsen, and Tranæs (2007); Rosholm and Vejlin (2010); Andersen, Dustmann, and Landersø (2019); Arendt (2020).

[11] Agersnap, Jensen, and Kleven (Forthcoming) use three reforms of the welfare benefit generosity to investigate the effect of welfare generosity on immigration rates, i.e. to test the "welfare magnet" hypothesis.

[12] In a more general context, Card, Kluve, and Weber (2010, 2018) argue that returns to job search assistance are small or zero, whereas returns to programs with a skill-training component are significant after a couple of years.

Section VIII contains a cost-benefit analysis and concludes the paper.

II Refugees in Denmark: Recent History

Following a period of low immigration, a significant number of refugees resettled in Denmark since the 1990s. Bosnians in the early 1990s, and Syrians in recent years, constituted the largest waves. In both cases more than 30,000 people were resettled in Denmark within a short period of time. In between these major influxes, refugee admissions to Denmark averaged 2,500 per year.

Displaced individuals who apply for asylum in Denmark are registered and placed in accommodation centers until their case has been adjudicated. These centers are usually in sparsely populated areas. Until 2013 the asylum seekers were not allowed to work while residing in these centers. Most asylum seekers, therefore, had very little contact with the Danish population prior to obtaining refugee status. In the late 1990s, the time of the considered reform, refugees waited, on average, more than one year before a decision on asylum was made (Table III). When the asylum status was granted, the refugee was resettled in a municipality chosen by placement officers in the Danish Refugee Council with the aim to distribute them across municipalities and encourage assimilation.[13] The municipality administration was responsible for finding public housing for the refugees and generally provided them with social support and unemployment services. Language training was offered up to a maximum of 1.5 years after resettlement by the Danish Refugee Council. Within this general background, a significant change in language training services occurred with the 1999 reform, as we describe below (section IV).

Similarly to the experience of other European countries, refugees in Denmark have shown a significant initial disadvantage and slow convergence to native earnings and employment probabilities. Refugees arrive without a job and are initially dependent on welfare. Their employment rate and their earnings grow during the first five to ten years after settlement. However, after that, they seem to stabilize at a considerable gap not only with natives but also with other immigrants.[14] The employment gap between refugee men and low skilled native-born men is 15-16 percentage points ten years after immigration (Schultz-Nielsen, 2017, Table 2). While the countries of origin have changed, the initial conditions and the difficulty in integrating refugees economically and socially have been persistent features of the Danish experience as well as of the experience of other European countries (Schultz-Nielsen, 2017; Bratsberg, Raaum, and Røed, 2017; Brell, Dustmann, and Preston, 2020). The next section de-

[13] UNHCR Quota refugees are resettled directly from refugee camps abroad. They constitute 10 percent of our sample.

[14] Appendix Figures 1.1 to 1.3 in Schultz-Nielsen (2017) show employment rates, earnings and welfare dependency of the refugee cohort analyzed in this study and two subsequent cohorts.

scribes our sample and the outcomes we analyze in detail.

III Data

Our sample includes all refugees and family-members reunified with existing refugees who were admitted in a four year window around January 1st, 1999. We restrict our attention to individuals who were between 18 and 49 years old at the time they were granted asylum and we follow them until 2016, spanning 18 years after the reform. The age range is chosen to include individuals who were old enough to qualify for the program at the time of arrival (i.e. 18 or older), and still young enough to be of working age (65 or younger) by the end of the considered period.[15] We include, as family-reunified individuals, the members of the immediate family (spouse and children) of a refugee in Denmark. These steps leave us with a sample of 8,558 individuals.

These 8,558 refugees can be linked to 16,598 children, 45 percent of whom (7,436) were born before arrival in Denmark and 14 percent (2,294) were born after arrival but before 2003. In order to observe schooling and social outcomes for these children we restrict the "*children sample*" to those 9,730 born between 1986 and 2002. This allows us to observe schooling outcomes into secondary school and outcomes in teenage years for most of these children. In particular, we observe graduation from lower primary school for everyone who is in the correct grade for age, and we also observe early measures of juvenile crime for the 16+ year old.

We use administrative data from multiple sources. The Integrated Database for Labor Market Research ("IDA"), the Income Register, databases on enrollment in and completion of schooling, and a database on charges and convictions for criminal activities are merged and used to construct the main outcomes.[16] We link these longitudinal data to the Population Register ("BEF"), the Admission Register ("OPHG"), the Migration Register ("VNDS"), and the register on internal mobility in Denmark ("FLYT"). The combined data provide key demographic variables, such as age and gender, identification of refugee status, date of receipt of refugee status, the country of origin, and the date of settlement in a Danish municipality. They also allow for merging the refugees with their spouses and children. Finally, we link each individual to the Register for Labor Market Policy Measures ("AMFORA"), which includes the services received in the municipality of settlement, and to additional information on the refugees

[15]We exclude a few individuals for whom no address could be found within the first two years of arrival, and we also exclude Somali refugees due to irregularities in the processing of asylum applications for this particular group in the last months of 1998.

[16]The Integrated Database for Labor Market Research (IDA) that we rely on to construct labor market outcomes such as employment and occupation is available from Statistics Denmark with a few year lag. This is the reason why we only follow the refugees until 2016 while we can follow the educational and crime outcomes of their children until 2018.

from the language training facilities ("Danskuddannelsesdatabasen"), such as their mother tongue and their initial assignment to language training track 1, 2 or 3, which is a proxy of their education and competences at arrival. Individuals with no or only basic schooling and individuals who need a thorough introduction to the Latin alphabet are usually assigned to Danish 1. Danish 3 is a fast learner track, and individuals with a long education from abroad or individuals who likely pursue further education in Denmark are usually assigned to this track, while Danish 2 is an intermediate track.

The summary statistics of the main outcome variables are shown in Table I. We measure employment as the fraction of a full-time working year. This measure takes the value one if the worker was a full time-employee during the whole year while it is less than one if the individual was either a part-time employee or not employed at some point throughout the year. Notice that the average annual employment in the sample of refugees was only 0.2 (20 percent) of a full working year, implying substantial non-employment or under-employment for this group throughout the sample period.

A second outcome we analyze is annual gross earnings measured in US Dollars and deflated to year 2000 prices as the base year. If no information on earnings or employment was provided in the registers, we assigned no employment and zero income to the observation. A considerable share of refugees had no earned income and so the average annual income is a very modest $8,600. In the empirical analysis we report the impact of the reform on outcomes in levels and we then translate them into percentage effects (relative to the pre-reform mean at the cutoff point).

We analyze a number of additional outcomes, to shed light on mechanisms behind the labor market effects and to have a more complete picture of the intended and unintended consequences of the policy change. These outcomes include additional education, occupational change, changes in municipality of residence and criminal activities. Table I shows that a large share of refugees was employed in Personal and Protective Services ($ISCO = 51$), Sales and Services ($ISCO = 91$), and as Machine Operators and Assemblers ($ISCO = 82$). These three occupations represent 83 percent of employment in the refugee sample and they are the top three occupations, not only overall, but also if we split by gender, age or language training track. Personal care ($ISCO = 513$) is the most frequent occupation for women, young individuals and those assigned to Danish 3. Moreover, the typical additional education pursued by refugees is as personal care assistants ("SOSU medhjælper" in Danish). Hence, many refugees assist health care professionals and provide personal care for elderly or persons with an illness or disability.

We also analyze an index of communication relative to manual intensity of the occupation of the refugees. This is calculated as the difference between the natural logarithm of the communication task intensity and the natural logarithm of the manual task intensity of the occupations, obtained by linking the occupations to ONET data.[17] Additional education is measured by enrollment in and completion of some education program in Denmark. 13 percent of the considered refugees have completed some additional education in the 18 years after settlement.

In terms of internal mobility, Table I shows that 57 percent of refugees were not living, by year 2016, in the municipality where they first settled. As for criminal outcomes, we use data on the number and type of criminal charges and convictions. In the main analysis, we focus on shoplifting. Property crime was the most common criminal activity in our sample, and the effects we find are driven almost exclusively by shoplifting crimes. The annual average crime rate was three percent in our sample with one percent represented by shoplifting crimes. The probability of ever being convicted of a crime for individuals in our sample was 24 percent (not reported in the table). Among those, property crimes were the more common category, with 17 percent of the sample convicted at some point of such a crime, and with shoplifting as the most common single crime (eight percent of refugees convicted at some point), as also pointed out in Andersen, Dustmann, and Landersø (2019). As a comparison, 13.7 percent of natives aged 18-49 in 1999 were ever convicted of a crime.[18]

Finally, for the children of refugees born before 2003, we analyze two types of outcomes. The first is a group of schooling outcomes; specifically, exams at the end of lower secondary school, graduation from lower secondary school, and enrollment in upper secondary education. The second is criminal outcomes, namely being charged with and convicted of some crime as a juvenile.

IV The Reform

The reform we consider is the first Act on Integration and the new Act on Language Education, which went into effect on January 1, 1999. Individuals granted refugee status before this date were subject to the old rules whereas individuals granted on or after this date were subject to the new rules. The purpose of the reform was to promote integration into civil society and improve labor market outcomes for newly arrived refugees. The measure to achieve these goals was a more extensive language training. Participation in the new program for people who qualified was mandatory as non-participants could be financially sanctioned. More importantly, acquiring permanent residence status required proof of

[17]The task outcomes use information from the O*NET database (US Bureau of Labor Statistics) merged to Danish register data via the International Standard Classification of Occupation.

[18]Notice that the conviction rates for this group of refugees decreased steadily over time, so that the early years were those with highest conviction rates for them (around 5 percent), while after 10 years in the country, the yearly conviction rate was less than two percent.

completion of the program.[19]

An additional temporary provision of the reform, aimed at increasing work-incentives for a subgroup of workers, was a cut in the welfare benefits. While the enhanced language training was mandatory for all refugees 18 years or older, the cut in welfare benefits was only for those older than 25 or with children. Welfare benefits to this group were reduced by 29 percent (25 percent if children present). Several studies, e.g. Andersen, Dustmann, and Landersø (2019) and Rosholm and Vejlin (2010), have analyzed the effects of a different cut to welfare benefits of immigrants, implemented in 2002. That cut was larger (around 50% of the benefits) and lasted for longer (at least seven years for newly arrived immigrants). The cited studies find significant positive employment effects of that cut. The cut we are considering was instead withdrawn after 13 months, on February 1, 2000. We will therefore separately analyze two groups: those below 25 years of age without children experiencing only the expanded language training and the rest of refugees, over 25, subject to expanded language training and the one-year welfare cut. This provides some insight into the potential additional work-incentive effect generated by lowering welfare benefits on a group.

The identification of the effect of this additional treatment is not as clean as the primary treatment (language training), as the treatment is associated to a demographic feature (age and having children). Heterogeneity across ages and family structure in the response to the primary treatment may be confounded with the impact of this policy. Additionally, there is some doubt about whether the policy of welfare reduction was fully enforced. A report released in December 1999 showed that affected individuals had been compensated by increases in supplementary transfers, potentially reducing the actual "size" of this treatment (Ministry of Interior Affairs, 1999).[20] Assessing the impact of expanded language training on short and long-run outcomes of refugees is therefore the primary goal of this paper, while evaluating the impact of the welfare cut is an additional secondary goal.

The new language training policy increased the instruction time from 1,370 to 1,800 hours and allowed such training to occur over an extended period of 36 months, doubling the 18-month limit existing before.[21] The structure and quality of language training was also changed; centralized goals

[19]Refugees only hold a temporary residence permit, and therefore, if not qualifying for permanent residency, risk being returned to their home country.

[20]The welfare portion of the reform was heavily debated already before it came into effect, and UNHCR stated in the summer of 1998 that the welfare benefit reduction "violates Article 23 of the UN Refugee Convention" (equal public relief and assistance for refugees and nationals). The UNHCR, however, agreed to the Danish Ministry of Interior Affairs' request to await an evaluation. In November 1999, the Danish Prime Minister announced that the reduced benefit did not work as intended, based on results from an evaluation published in December 1999 (Ministry of Interior Affairs, 1999). The evaluation showed that very few immigrants could supplement the welfare benefit with income from work, which was part of the motivation for the reduced welfare benefits.

[21]The reform resulted in more extensive language training than offered to immigrants in most other countries (e.g. the German introductory language course is 600-900 hours and Sweden offers on average 525 hours of language training).

and national tests were introduced and resources to increase the qualifications of the teachers were provided. The Act on Integration also introduced a course in civic education. While important, this was more limited in scope when compared to the great boost of language education. Only 20 hours were dedicated to civic education, compared to 430 hours of additional language training. In the rest of the paper we refer to this whole package of augmented language instruction and civic orientation as *"language training"*. The active labor market policies (ALMPs), including on-the-job training, provided to refugees were not changed in any way by the reform and we refer to them as *"employment support"*.

A minor administrative change introduced by the reform was that, after January 1, 1999, the municipalities (rather than the Danish Refugee Council) became responsible for managing language training. They also remained responsible for managing the ALMPs for all residents.[22] Therefore, the reform combined the responsibility for language and employment services in one place, namely the municipalities. As it took around two months between obtaining refugee status and resettlement in a Danish municipality, refugees arriving within 2-3 months of the cutoff date received their employment and language support from the municipalities. This implies that, at least when we consider rather small bandwidths around the cut-off, treated and untreated refugees both experienced language training administered by municipal government.

To ensure that municipalities would provide the enhanced language training to refugees who qualified, the reimbursement per refugee was larger for those arriving after January 1, 1999 relative to those who arrived before. The reimbursement was constituted by a fixed monthly payment per refugee and activity-based payments for the months of employment or language services. Both the fixed and the activity-based payments were higher for refugees arriving after the reform and they lasted twice as long (36 compared to 18 months).[23] Hence, the financial incentives of the municipalities were also aligned to provide the expanded services only to "treated" refugees, i.e. those arrived on or after January 1st 1999.

Table II shows the estimates of discontinuity, at the reform cut-off date, in the participation to and intensity of language training and employment support (which should not be affected by the reform). The top two panels (a and b) show the regression discontinuity (RD) estimate of the share who ever participated to the language program and to the employment support program, respectively.[24] Below the

[22]The Danish Refugee Council is a private non-profit organization. The Danish state finances language training for the population we study both before and after the reform but the administrative unit that refers refugees to language schools changed.

[23]The fixed payments were $340 and $515 per month, respectively, for refugees arriving before and after the reform. The activity-based payments were $212 and $288 for employment services and $500 for language services to refugees arriving after January 1, 1999, whereas no earmarked payment was made for language services to refugees arriving before the reform, as per Act on Integration 1999, section 45(5), 45(6) and 45(7) and section 59(3) and 59(4).

[24]The data refer to participation in January 1999 or later as 1999 is the first year we observe the intermediate outcomes shown in Table **??**. We use the specification in equation (1) for all RD estimates, allowing for different slopes at either side of

estimates, we also show the mean of the participation share for the control group at the cut-off. Different columns show the effect on share from one to three years after the reform. None of the estimates is significant at the five percent level, implying that participation into the language program did not really change with the reform. Nor did participation into the employment support program. Panel c shows, however, that what substantially changed for the treated group was the hours of attendance to language training. Qualifying refugees accumulated roughly an extra 200 hours of language training, on average, relative to non-qualifying ones. This is about one fifth more language training than the control group (980 hours at the boundary point). While very large, the additional training hours logged by the"treated" group, on average, was less than the total potential extra hours (total of 430) for which they would qualify. Additional data from the training facilities reveal that treated and control refugees were rarely in the same classes, which suggests that language support was possibly of different quality. Panel d shows that the cumulative hours of active labor market support were unaffected by the reform, confirming that this reform did not change other services provided to refugees.

The new Act on Integration also modified few other aspects of the refugee resettlement policy which had been in place since 1986. First, the welfare benefits were made conditional on residing in the assigned municipality for the three-year Integration Program, and on participation in the program.[25] This made it more likely for refugees to comply with the training requirements. Refugees subject to the new rules were more likely to stay in the assigned municipality until completion of the program and postponed mobility, as we show in section VI.B, Figure VII.

Second, the geographic dispersion of refugees was decided in a more mechanical way. While before the reform bilateral negotiations between the municipality and the central government determined the number allocated to each area, after the reform refugees were distributed according to county quotas decided by the Danish Immigration Service ("Udlændingestyrelsen"), taking the number of immigrants already living in the municipalities into account. The municipalities within a county were then supposed to agree upon the within-county settlement, but if they failed to agree, the central government would determine municipal quotas. This formal change in procedure does not imply any specific change in sorting or selection in the distribution of refugees across municipalities. To rule out that it generated different sorting, potentially confounding the impact of increased language training, we include a set of municipality characteristics in the balancing tests between treated and control refugees that we conduct in section V, Table III. We find no differences in receiving municipality characteristics between

the cutoff which arise mechanically only in Table II due to the lack of these data prior to 1999.

[25]Refugees subject to the new Act on Integration can only move municipality of residence and keep their benefits provided the new municipality agrees to take over the responsibility of the continued immigration process.

treated and control refugees. In particular, the share of co-nationals and the economic performance were essentially identical for host municipalities of treated and non-treated refugees.

V Identification and Empirical Design

V.A Empirical Specification

Given the sharp discontinuity on January 1st 1999, to the eligibility for the enhanced language training, we apply a regression discontinuity design comparing refugees admitted immediately before and after the discontinuity date. In formal notation, the treatment variable for individual i, D_i, is a deterministic function of the date the refugee status was granted x_i:

$$D_i = 1\{x_i \geq c\}$$

where c is the cutoff date of January 1, 1999. Hence, $x_i - c$ is the time in days between the cutoff date and the date of admission and it is our so called "running variable". In our baseline specification we assume linear trends on either side of the threshold date, so that the estimated equation is as follows:

$$Y_{it} = \alpha + \tau D_i + \beta_1(x_i - c) + \beta_2 D_i(x_i - c) + \varepsilon_{it} \tag{1}$$

The coefficient τ is the regression discontinuity (RD) estimate and captures the causal average treatment effect of the reform on outcome Y_{it}. The two linear terms $\beta_1(x_i - c)$ and $\beta_2 D_i(x_i - c)$ capture the linear dependence of outcomes on admission dates measured relative to the cut-off date.[26]

The variable Y_{it} in equation 1 represents the outcome of interest measured at $t = 1, 2...18$ which represents years after the reform. We estimate a separate regression for outcomes in each year from 1 to 18 after the reform using Weighted Least Squares with a triangular kernel to give more weight to the observations closer to the cutoff. We restrict the observations used in the estimations to a narrow window around the cutoff, determined by the optimal bandwidth selection algorithm from Calonico et al. (2019). This mean squared error (MSE) optimal bandwidth varies across outcomes and years. It ranges between four and nine months around the cut-off for the main outcomes analyzed and is never beyond one year. We also show that the point estimates are not very sensitive to the exact choice of bandwidth, nor to

[26]We analyze specifications with other functional forms including second and third order polynomials in $(x_i - c)$ in the robustness checks. The main results are insensitive to the choice of functional form of the running variable. The recent consensus, exemplified in Gelman and Imbens (2019), is in favor of the local linear specification.

functional form assumptions.[27] We use heteroskedastic-robust standard errors (as suggested e.g. in Lee and Lemieux, 2010).

It is important to note that we obtain a reduced form effect (intent to treat effect) of the reform and not the average treatment effect on the treated of the 430 additional hours of eligibility to language training, since not all participants attended the maximum number of hours (in fact we showed in Table II above that on average they attended 200-250 extra hours). For the subgroup of 25 years and older we capture the additional average effect of a cut in welfare benefits. Below we discuss the validity of the identifying assumptions.

V.B Density of Cases and Balancing Tests

Asylum seekers obtaining refugee status in Denmark around the cut-off date applied for asylum long before the new legislation came into effect. Most refugees, at the time of applying, had no knowledge of the changes that were about to happen. The Act on Integration was proposed on April 16, 1998, and passed on June 26, 1998. The average waiting time for decisions on applications for asylum was 13 months at the time (Table III), and the median was slightly more than ten months. Hence, a large majority of the refugees who were granted asylum near January 1, 1999 had applied before the law was even proposed.[28] The Danish Immigration Service is in charge of granting residence permits, and the exact admission date is outside the control of refugees as well as of the municipalities, which provide the integration services after resettlement. We are not aware of any change in the processing time or in the order case workers in the Immigration Service processed asylum applications. To test this formally, we perform two types of checks that strategic behavior or timing in approving refugee cases did not occur.

First, we test that the number of approved applications does not show any unusual bunching (increase or decrease) on either side of the cutoff. Figure I shows the raw distribution of admissions, as share of total admissions over the the period, by month from January 1, 1997 to February 28, 2001 in Panel a. Then, to account for seasonality and time trends in refugee admissions, we show, for each month, the share of annual admissions in deviation from the mean share for that month over the four years in Panel b.[29] A value of zero in the observations of Panel b implies that the admissions of the month equal the average for that month over the four years shown. Very little variation both in the raw count (Panel a)

[27] For outcomes such as employment, that is marginally significant with the MSE-optimal bandwidth, statistic significance at the five percent level does depend on bandwidth.

[28] See Hvidtfeldt et al. (2018) for more information on the process and waiting times in the Danish asylum system.

[29] The admission data are only available from January 1997. Our full sample includes refugees arriving within a four year window around the reform as explained in section III. Figure I includes two additional months to better illustrate the seasonality around New Year.

and in the "de-seasonalized count" (Panel b) is observed from December 1998 to January 1999 (Panel a). January is right at the mean for this month and December is slightly below (-0.03 percentage points). The most unusual months (out of the 50 months we show) are April 2000 (-0.05), April 1997 (0.04) and May 1998 (0.04). We conclude that the density of applications is smooth around the cut-off date and nothing indicates that the approval process was slowed or accelerated around the cutoff.[30]

A second possibility is that case workers continued to process a uniform number of cases around the cutoff but were selective about which cases to allow into the new policy regime. Moreover, the small procedural change in the resettlement process, starting on January 1, 1999, might have generated differential sorting of refugees across locations. Table III shows the summary statistics of predetermined individual characteristics in our sample (Panel a) and of the characteristics of the municipality of resettlement (Panel b). Column (1) and (2) show the sample mean and standard deviation of these variables. Then, column (3) shows the difference around the RD threshold for the individual variables and for characteristics of the assigned municipality. They are estimated using equation (1) and column (4) shows the confidence interval for testing the hypothesis that the discontinuity is equal to 0. These are checks that the distribution of individual characteristics and of first assigned municipalities are random around the discontinuity.[31]

The average age of refugees at admission in the sample was 31.5 years (after excluding individuals older than 49 as we follow individuals for 18 years). The typical refugee was married (67 percent), male (58 percent) and the average number of children was slightly above one at the time of admission. We also report the percentage of refugees admitted under different channels, such as the Geneva convention, UNHCR quota refugees and family-reunification with existing refugees. The majority, 53 percent, are granted refugee status under supplementary Danish laws. The single most important country of origin around the reform was Iraq, representing 44 percent of our sample.[32] 16 percent of refugees were from Afghanistan, and the rest was from a range of smaller origin countries (see Section III for a thorough description of our sample and data).

The Table (columns 3 and 4) shows no evidence of differences in composition and characteristics of refugees on the two sides of January 1, 1999. They are of similar age, equally likely to be married and have the same number of children. The refugees could not be employed upon arrival and we do not have

[30]Besides evidence of no manipulation, even if refugees could manipulate the date, it was not clear ex-ante that they would want to qualify for the new policy. While it increased language training, it was combined with sanctions for non-participation and a reduction in the level of welfare benefits.

[31]We use local linear regressions with the triangular kernel and the optimal bandwidth selector from Calonico et al. (2019) in order to test the discontinuity. Appendix Figure A.I illustrates the balancing tests for each of the variables in the Table I.

[32]Iraqi refugees in our sample fled Saddam Hussain's regime before the Iraq War started in 2003.

information on education received in their country of origin. Instead, we use the assignment to language training as a rough indicator of skills. The language training facilities place refugees into Danish 1, 2 or 3 based on their skills at arrival. Those with primary schooling or less are typically assigned to the lowest level, Danish 1. Tertiary educated refugees are most often assigned to Danish 3, and Danish 2 is an intermediate group. Using this measure of educational attainment, the tests in Table III show no differences in skills around the cutoff. Finally, we see from the summary statistics that 35 percent of the admitted refugees spoke Arabic as their mother tongue, while 9 percent reported Dari as their first language. These are the most common languages in Iraq and Afghanistan.

In terms of their geographic distribution, 33 percent of refugees were placed in urban municipalities (Table III Panel b). The RD estimate is 0.11. Although not significant, this is a relatively large difference. However, it turns out that it is driven by municipalities on the margin of being classified as urban. Despite the fact that 17 percent of our sample is placed in one of the five largest cities in Denmark, there is no difference across January 1, 1999 in this variable (RD estimate 0.01 and standard error 0.05).[33] So there is no evidence of selective placement in the largest cities. Additionally, the treated and the control groups are distributed in municipalities with no significant differences in employment and unemployment rates, in average labor income and also in number and share of co-nationals already residing there. Overall, the balance test show a remarkable similarity of observable characteristics of refugees and of the locations where their were settled, before and after the cut-off date.

VI Estimated Effects on Refugees

VI.A Main Labor Market Effects

Figures II and III show the main results for labor market outcomes. They represent the RD point estimates and the 95-percent confidence intervals of the impact of the reform on the employment rate and on gross earnings of refugees, respectively, for each year from 1999 (year 1) to 2016 (year 18). The two top panels (a and b) show the average estimates for the whole group of refugees, while the two bottom panels (c and d) show separately the effects every fourth year for refugees aged 18-24 with no dependents (children), who were only subject to the increased language treatment, and those above 25 who additionally had a reduction in benefits in year one (as explained in section IV).[34] Panel a and c

[33]The five largest cities are Copenhagen (incl. Frederiksberg), Aarhus, Aalborg, Odense and Esbjerg.

[34]We exclude from the 18-24 year old group 518 individuals (30.6 percent) with children in the household because all individuals with a child experienced a reduction in benefits to 75 percent of the previous level. To keep the group of 18-24 as untreated by the reduction in welfare, we include individuals of age 18-24 with children in the other group. Effects by finer age groups are shown in Appendix Figure A.II.

plot the estimated impact in each year, while Panel b and d plot the average cumulative impact, on the employment rate (in Figure II) or on earnings (in Figure III).

Looking first at the average effects over time, in Panels a and b, two features emerge clearly for employment (Figure II). First, while we see small to no labor market effects on treated in years 1 to 4, when the refugees were likely attending the language classes, we see positive effects from year five on, when the classes were completed. The magnitude of the employment effect is in the order of 7-8 percentage points of a full working year in years 6 to 8 after the reform. We also observe a decline in the effect in year 9, 10, corresponding to the recession of 2009-10 and then stabilization around 6-7 percentage points. The second fact emerging from the graphs is that the long-term (average cumulative) effect stabilizes around 4 percentage points per year with a small dip during the recession.[35]

The profile and time pattern of the treatment effects on earnings (Figure III) are very similar to those on employment, with annual earnings approximately $3,100 higher for the treated group 6 to 8 years after the reform. As labor income is the main source of earnings of refugees, this is expected. These are large percent effects due to the low baseline of employment among refugees. From 6 to 7 years after the reform, annual employment was about 38-56 percent higher for the treated group and annual earnings about 42-50 percent higher. Over 18 years this accumulated into a total of 0.76 years (9 months) of extra employment, corresponding to a 4 percentage point higher employment rate in each year for the treated. In terms of earnings, the treated group accumulated $45,267 higher income from employment (measured in 2000-prices) over 18 years, which corresponded to an average of $2,510 higher earnings per year. Relative to the baseline these are large effects, equal to 23.0 and 33.7 percent rise in employment and earnings, respectively.[36]

The precisely estimated zero difference in employment and earnings between the treated and the non-treated during the first 3-4 years is consistent with a labor market effect of language skill acquisition. During the first 3 to 4 years refugees were learning and the new skills generated an impact on labor market outcomes in the following years.

Panels c and d of Figures II and III separate individuals who were aged 18-24 and without children in 1999, as they had no change in welfare in the first year while the others experienced a drop in welfare for one year. Panels c shows the effects every four years for the younger group (in blue) and the older

[35]Employment is measured as a fraction of a full working year. Hence, annual outcomes can be interpreted as a fraction of one year full-time equivalents and cumulative outcomes may be interpreted as a fraction of years of employment. Effects are more noisy if we split employment into intensive and extensive margins, but we similarly find relatively small coefficients in the first couple of years (more on this below, in Figure A.IV).

[36]The baseline is 0.15-0.19 average years worked (annual) and $6,159-7,371 average yearly salary (annual) 6-7 years after the reform, and 3.3 worked years (cumulative) and $134,378 (cumulative) by 18 years after the reform for employment and earnings, respectively.

group (in red) separately. Panel d shows a similar figure with effects accumulated over every four-year period. While the effect on the young is estimated with a substantial noise due to the small sample, there are no significant differences in the effects between the two groups. The average accumulated effects, in particular, are very similar for the two groups, both for employment and earnings. If anything, in the yearly effects we see a positive but not statistically significant effect for the young group in year one, while the other group shows a very precise zero effect. This does not support the idea that the one-year reduction of benefits generated a work-incentive effect on the older group subject to welfare cut. Similarly, the effect on earnings is larger (but not significantly) in the first years for the 18-24 years old and equal for the two groups in the later years. The cut in benefits did not produce more employment or more learning, consistent with the finding in the report by the Ministry of Interior Affairs (1999) that led to the abandonment of the benefit reduction. This is in line with the idea that the early hurdle to employment for newly arrived refugees was a lack of skills, language among them, affecting their employability rather than a lack of willingness to be employed. Overall, the effects for the two age groups are similar and we can never rule out homogeneous effects on earnings and employment for the two groups or even across a finer partition of workers into six-year age groups (see Appendix Figure A.II). While it is possible that the zero employment effect was due to the fact that the welfare cut was not large enough, or to the fact that our sample size is not large enough to precisely identify small effects, the evidence from these results is also consistent with the idea that there is no quick fix to increase employment of refugees in the early years. However, improving their language skills through increased training pays off after 4-5 years when they can use those skills in the labor market.

Figure IV presents a series of robustness checks for the main results, synthesized by the coefficient on the average cumulative employment and earnings 18 years after the reform. This coefficient captures whether there is a persistent long-run effect of the reform and its magnitude in average employment rate and earnings per year. The first two panels (a and b) show that the estimated magnitude of the effects are very robust to the choice of bandwidth. From bandwidths of 60 to 360 days (including the optimal bandwidth which is 216 for employment and 277 for earnings) the obtained estimate is very stable and, in most cases, statistical significant.

In Panels c and d we show the estimates obtained at arbitrary cutoff dates before and after the actual cutoff date. The red line shows the estimate of the impact at the actual discontinuity. In most cases at arbitrary cutoff points we estimate very small and non-significant effects. Occasionally some significant effects are estimated. Since the reform was implemented on January 1, 1999, it is also reassuring that there is no effect of January 1998 (-365) or January 2000 (+365), ruling out that some reoccurring seasonal events related to the new year are driving our results (three out of four point estimates are zero and the last one has the opposite sign to our results).

In Panels e and f we see that the estimates are robust to a series of additional checks. First we include control variables in the regression, then we use second or third order polynomials instead of local linear regressions. In the "donut" specification we exclude refugees admitted in the two year window around January 1, 1999. Finally, we also show the estimates from a simple OLS regression where the time to cutoff is a linear variable and obtain a similar estimate of the discontinuity effect. One can also read the stability of the estimated long-run effect with respect to bandwidth as well as the "donut" estimates (omitting one year on each side) as showing that the impact on those arriving right after January 1999, who were exposed to a full year of welfare benefits reduction, is not different from the impact on those arriving after one year and who experienced almost no welfare reduction at all (as the benefit reduction was discontinued in February 2000). This is another piece of evidence that is consistent with no employment effect from the welfare cuts.[37] Overall, the results are rather insensitive to the inclusion of controls and to the change in bandwidth, suggesting that the cutoff was indeed random. Using other functional forms and a simple OLS estimator one would also get a similar, although slightly larger, estimate (Panels e and f).

We also investigate (see A.IV in the Appendix) the margins of the employment and of the earnings responses. There is no evidence that hours of employment (the intensive margin, measured as a fraction of the working year) varies with treatment for employed individuals (Panel a of A.IV in the Appendix). Similarly we find no effect on the share working full time, conditional on working (Panel b of the same figure). Turning instead to extensive margin responses, we find a tendency for the treated group to have higher probability of working, starting 5 to 6 years after the beginning of the treatment period (Panel c). The annual effects are rarely significant. If we instead measure the effect on having ever been employed, treated refugees are 7.8 percentage points (not statistically different from zero) more likely to have had a job by the end of the analysis window. We also find that earnings conditional on being employed rise (Panel d), revealing that earnings effects are not driven solely by higher employment but also by higher earnings among the employed.

[37] Appendix Figure A.III shows the sample means of the outcomes 18 years after the reform by one-month bins around the cutoff and a linearly fitted regression line. These graphs confirm that there is a discontinuity in the outcomes around the reform and clear level difference before and after.

VI.B Effects on Schooling, Occupational Change and Geographical Mobility

We now show, in Figures V, VI and VII, the effects on outcomes that can be potential mechanisms helping to explain the observed labor market outcomes and their dynamics.

Panels a and b of Figure V show that treated refugees were more likely to pursue additional education in Denmark. The average effect emerges 8 to 9 years after the beginning of treatment, i.e. 4 to 5 years after the end of the language training. The timing of the increase in education coincides roughly with the aftermath of the great recession. When separating the young (under 25) from the older in Panels c and d, we see that the effect is driven exclusively by the young. This group is more likely to obtain lower secondary education (significant coefficient) and a subgroup of them may also obtain academic education in Denmark (positive but non significant coefficient in Appendix Figure A.V, Panels b and f). The largest effect on skill upgrading is represented by attending vocational education programs. These programs lasts for 2 to 4 years and can be attended while working (Appendix Figure A.V, Panel d).[38] Interestingly, the effect on attending vocational classes emerges later and is significant from around 9 years after 1999. This is consistent with a scenario where refugees who spoke better Danish due to extra language training were more likely to choose to improve their vocational training if they lost their job during the recession, to consolidate their strengths in the Danish labor market (Appendix Figure A.V, Panels a, c and e show enrollment). This transition to more education is what likely allowed them to obtain better paid jobs after The Great Recession.

Figure VI look at occupational change conditional on being employed (in November). Panels a and c show that treated refugees tended to upgrade their occupations towards jobs with higher communication (and less manual) content. The transition towards more communication intensive jobs is especially evident 4-5 years after 1999, when refugees likely completed their language training and obtained their certification. The effect is still significant in the medium run but it seems to diminish after the economic crisis and it is not significant in the long-run. Panel c shows that both young and old refugees experienced this occupational upgrading in the early years after completion of the training, with a potentially stronger long-run effect for older refugees. While the younger group may have mainly benefited from better language training by adding some formal vocational schooling (especially after the crisis), the older group simply transitioned to less manual-task intensive occupations. Panel b and d show a similar story with significant effects on the share working as Personal and Protective Workers in the medium run. Learning better Danish allowed many of the refugees to be employed in occupations like nurses, health

[38] Paid work-based learning is part of the vocational training programs.

care assistant or personal assistant for elderly citizens. [39]

Finally, in Figure VII we analyze whether the reform affected refugees' geographical mobility. Panels a and c show the impact on the probability of having moved from (not living in) the municipality of initial resettlement each year after treatment. As usual, Panel a shows the average yearly effects and Panel c shows the effect by age group every four years. The pattern is very interesting, and it reveals a lower probability that treated refugees relocate out of the municipality in the early years, but then no difference in the probability of relocation in the long-run. Such a pattern reveals a delay in the probability of moving from the municipality of initial settlement. Individuals qualifying for the additional language training were more likely to stay in the first 3-4 years after 1999 (likely to take advantage of the services) and to leave in the following years so that their likelihood to have left was equal after 6-7 years. Given a significant rate of overall mobility of refugees from their initial municipality of resettlement, this shift in time is aligned with the incentives of the treated group to stay for the years of language training, and then leave, potentially to pursue jobs, further education or a location with better amenities for them. The lower mobility effects appear a bit stronger for the young group who may have stayed longer in the original municipality to find a first job, while the older, who also experienced a reduction in welfare, were more likely to move for instance to stay with relatives. Panels b and d reveal no differential tendency of treated refugees to be in or move to urban areas, neither soon after the treatment nor in the long run.

VI.C Effects on Adult Crime

Labor market success and economic integration are desirable and important goals for refugees. Achieving them would improve their material and psychological well-being, would increase local benefits from hosting them and may improve the opinion of natives about them (Bansak, Hainmueller, and Hangartner, 2016). In terms of affecting the opinion of citizens, however, there is an even more important dimension of integration which relates to potential criminal activities of refugees. Either as a consequence of their marginalization, trauma or disenfranchisement, there is a diffused perception that refugees are more likely to commit crimes. Statistics from several European countries usually show higher crime rates for refugees than for natives.[40] For Denmark, Andersen, Dustmann, and Landersø (2019) report that the refugee crime rate is about double that of similar natives, and most of this difference is due to property crime, mainly shoplifting. Our summary statistics broadly confirm these averages for the group

[39]We do not see a clear pattern in the two other top-three occupations discussed in section III and effects are not shown in the paper.

[40]Several German politicians and newspapers linked an increase in crime in 2015 to the arrival of an exceptionally high number of refugees and asylum seekers (see BBC News, 2018).

of refugees considered in this paper.

Therefore, it is important to analyze whether the language training and/or the reduced welfare treatment analyzed in this paper had any impact on the probability that refugees commit crime. Figure VIII shows the impact on the average annual probability of being charged with (Panel a) or convicted of (Panel b) a shoplifting crime in supermarkets. Panels b and d show the same figures, divided between under 25 with no dependents and others, every four years. We focused on this crime because shoplifting is the most common crime and it drives the effects measured on overall crime rates (see Figure A.VI). Panels (a) and (b) show that the probability of being, respectively, charged and convicted with a shoplifting crime is eight and seven percentage points higher for the treated group in the first year after the reform. Already in year two, however, and in all following years, the difference between the treated and non-treated group becomes very small and insignificant. As the first year after 1999 is the period in which the subgroup over 25 years experienced reduced welfare benefits, it is interesting to analyze whether this is also the group experiencing a higher increase in criminal convictions. This is exactly what is shown by Panels c and d. The older group of treated refugees shows a higher propensity to be charged and convicted in the first year, but in all the subsequent periods there are no significant difference in probability of being charged and convicted between treated and control. These findings do not support the idea that language training has any short or long run effect on the probability of refugees committing crimes. However, they are suggestive that the one-year reduction in welfare payment was associated with higher shoplifting by the group subject to that measure. While the coefficient is precisely estimated, and the contrast with the younger group which was not subject to the cut is striking, we do not want to overemphasize this result. The two groups may differ on other dimensions and the period in which this policy was in place was short. Nevertheless, this finding is consistent with what Andersen, Dustmann, and Landersø (2019) find in response to cut in welfare benefits to immigrants in the later reform of 2002.

VI.D Heterogeneous Effects

Table IV shows the RD estimates separately by gender (Panel a) and by alphabet of the mother tongue of the refugee (Panel b) for five different outcomes, each described by the column header. Panel a shows that the reform had a similar impact on men and women. Neither for labor market outcomes (columns 1 and 2) nor for education (column 3) the differences in point estimates between genders are statistical significant. Similarly, when we look at crime rates in the first year after the reform, the difference between men and women is not statistically significant. Surprisingly, the point estimates of the treatment effect on crime are larger for women (mothers) than for men, suggesting that their propensity to commit shoplifting increased more than men's in response to the one year reduction in welfare income. This suggest a "poverty" motive, rather than criminal behavior, at the root of the shoplifting crime.

The Danish language course is comprehensive by international standards and trains speaking, listening, writing and reading. Some refugees can neither read or write in their mother tongue, and the first language of the majority of the refugees uses the Arabic alphabet. Learning Danish which is based on the Latin alphabet might therefore be particularly difficult and particularly important for them. To test this hypothesis, we create two broad categories based on the mother tongue of the refugee: languages written using the Arabic alphabet, and the residual consisting of Latin and Cyrillic based script and of individuals whose mother tongue was not reported. Those recorded in the data from the language training centers as speaking Arabic (35 %), Dari (9 %), Farsi (4 %), Pashto (3 %), Assyrian (1 %) and a number of smaller groups are categorized as Arabic alphabet. The largest groups within the category Latin, Cyrillic and unknown, are Unknown (17 %), Kurdish (8 %), Albanian (6 %), Serbian (5), Bosnian (3) as well as Tyrkish (1 %). Table IV provides suggestive evidence in favor of our hypothesis. Labor market effects are much larger for those who likely do not know the Latin alphabet upon arrival.[41] Interestingly, the short-run crime effects seems to be driven by the other ethnicities.[42]

VII Effects on Fertility and on the Second Generation

The statistically significant and economically important effects of language training on employment and earnings of refugees may generate consequences that propagate to their family life, their social integration and, possibly, to their children, the second generation of immigrants. First, as treated refugees are more likely to be employed and to receive education in Denmark, they have a more direct exposure to the Danish culture, and this may have an impact on their attitudes and culture, including marriage and fertility choices. Second, their better economic situation and higher degree of integration in the Danish society may have affected investment in their children and the outcomes of the second generation. We analyze these two aspects in turn.

[41]Consistent with this we also find larger effects on individuals assigned to Danish 1 (not shown here) which is designed for those with little schooling from their origin country and those who cannot read or write at arrival.

[42]Formally, testing for equal effects in a fully interacted model is demanding and the results are shown in Table A.I. The interaction terms between the subgroup and the treatment variable shows that effects are never statistically different.

VII.A Marriage and Fertility

In this section we analyze whether the language training eventually affected marriage outcomes and the probability of having children, their number and their timing (i.e. the age of the mother at birth). Greater assimilation into the Danish culture may modify preferences towards marrying outside of the ethnic group or having fewer children, as average fertility of native Danish women is significantly lower than that in the countries of origin of refugees.

We first look at inter-ethnic marriage, namely the probability of marrying a person (Danish or immigrant) of nationality different from their own. Intermarriage is often a strong sign of social integration. The RD estimates (available upon request) are quantitatively very small and do not show any significant effect of treatment. We then analyze several outcomes related to fertility. Figure IX shows four graphs illustrating the effect of treatment on four different indicators. Panel a shows the average number of children at completed fertility for refugees, grouped by monthly bins and spanning two years before and after the cutoff date of arrival. The panel shows that number of children born after admission is the same on the two sides of the cutoff and no discontinuity is visible in a simple local linear approximation at the cutoff. The control as well as the treatment group both have a little more than one child after arrival. We also find no systematic differences in the timing of child births after admission. The treated and non-treated group both have a similar probability of having a child in each year after 1999 (Panel b), whether the first or a later child (Panels c and d).[43] These insignificant effects on inter-marriage and on the fertility rates of the treated group suggest that the language training did not change in any significant way the structure, size and composition of the families of the refugees. This implies that effects on their children, if present, are likely arising from the better skills, economic conditions and education of their parents rather than from changed size or type of family. We analyze children's outcomes in the next section.

VII.B Children of Refugees

In order to study the effect of parents' treatment on children outcomes, we need to choose whether the treatment affects at least one parent, or specifically the father or the mother, or both. This choice will also determine who constitutes the control sample. As the father was admitted before the mother in 57.3 percent of the children sample, and in 19 percent of the cases the two were admitted together, choosing the treatment of the father implies, in two thirds of the cases, that both parents were treated. We consider

[43] The graphs look similar if we study fertility from 9 months after admission, so that we only capture fertility decisions after admission status is known.

the admission of the first parent (which implies that both parents were treated) as our main specification in column 1 and 5 of Table V. In this case the control group includes children with zero or one treated parent. We also consider, as an alternative, having the father treated and we show the results in column 2 and 6. Alternatively, we consider admission of the last parent as the assignment to treatment (column 3 and 7) or the treatment of the mother (column 4 and 8). Both of these cases imply that it is more likely that these children have only the mother treated while the father, who arrived earlier, is untreated. In measuring children's outcomes, in Table V we distinguish between boys (columns 1 to 4) and girls (columns 5 to 8).

We select all children born before 2003 as our sample, and we analyze five outcomes, following them up to 2018. These outcomes are: taking any exam at the end of the lower secondary school, completing lower secondary education, enrolling in upper secondary education and criminal charges and convictions. The youngest of the children considered in this sample turned 16 in 2018, while the oldest were thirty-two years old in 2018.

Children born after admission and before 2003 were typically born while their parents were attending the three-year language program and they would be of school age after their parents completed the program. Children born before admission, on the other hand, were on average 6 to 7 years old (based on the admission date of the first parent) when their parents were attending the three-year integration program.[44] The estimates of column 1 of Table 3 show that boys of treated parents were 16 percentage points more likely to reach the final exams in lower secondary education (9th grade graduation exams), 11 percentage points more likely to graduate and 16 percentage points less likely to be charged with a crime by 2018, relative to similar children of non-treated refugees. No significant effects on enrolling in upper secondary education are found. The last two rows show that the treatment of both parents, and specifically of the father, seems to be associated with significant crime reduction for boys. The estimates on crime are smaller and not significant when we consider as treated those children with only the second parent (or the mother) treated (see columns 3 and 4). Looking at columns 5 to 8 in the table, which show the coefficient on girls' outcomes, we find very small point estimates and no significant coefficient on any of the variables. A large body of literature emphasizes how boys' basic school outcomes may be more sensitive to the family conditions and especially to the presence and role played by the father (see for instance DiPrete and Buchmann, 2013). In our case, if the father of the boy had acquired better Danish language skills, was possibly more integrated in Danish society and more likely to value the

[44]Table A.II shows descriptive statistics for children born before 2003, split into children born after admission of the last parent in Panel a and children born before admission of the last parent in Panel b.

education of his children, boys may have been strongly affected by that. Figure A.VII in the Appendix shows a series of robustness checks for the main outcomes for which we find a significant effect on boys, namely attending exams in lower secondary school and being charged with a crime.[45] The point estimates are mostly statistically significant but vary in size across the robustness checks we perform. In general, Panels e and f of Figure A.VII show that methods that place lower weight on observations close to the cutoff and use a wider range of the data to estimate the impact produce smaller and more credible results. There are a few outlier observations near the cutoff that affect the estimates. Higher order polynomials, allowing a closer fit at the cutoff, produce even more extreme observations, while OLS and a “donut” approach to RD design (leaving out two weeks on either side of the cutoff) produce smaller estimates. In most cases the estimates are negative and significant for criminal charges for boys.

VIII Discussion, Cost-Benefit Analysis and Concluding Remarks

In this paper we analyze the impact that an expansion of language training for refugees in Denmark had on their employment, earnings, education and other important social outcomes. The introduction of this policy, applicable to refugees who obtained refugee status on or after January 1, 1999 allows a causal estimation of its effect, due to the random distribution of refugees around the cut-off date. After eighteen years, refugees who got the expanded language training were (permanently) about four percentage points more likely to be employed (a 23 percent rise relative to the baseline) and earned \$2,500 per year more (a 34 percent rise relative to the baseline) than the control group. These effects appear to be permanent as they are significant and stable after 18 years. It is informative to perform a cost-benefit analysis of the government’s investment in the enhanced language training.

We conduct a simple cost-benefit analysis by calculating the discounted stream of benefits from the policy minus the costs. We include three types of costs and benefits in the base scenario. The main benefit is the increase in annual earnings for refugees as described above. The main cost is the operating costs of the extended language courses. This is approximated by the monthly activity-based payment for language services during the extended time spent in the language course.[46] The language course lasted for about 18 months prior to the reform, and the reform raised the resources for the language course by 30 percent, so that the time spent in language training increased by 6 months, distributed during the second and the third year of the program. The third component in the cost-benefit analysis is the change

[45] Using being convicted or being charged with a crime as the outcome variable in the analysis show similar patterns. Similarly, we leave out girls in the robustness checks for brevity, as no significant effects are found for them.

[46] The activity-based payment for language training is found in Act on Integration No 474 of 1998, section 45(6).

in the deadweight cost of taxation. This is calculated as the impact of the reform on the fiscal budget multiplied by a tax distortion rate. The impact on the fiscal budget is the sum of the operating costs of the extended program, the effects on tax revenue and savings on welfare benefits. The savings on welfare benefits are estimated using individual data on welfare payments (results available upon request), and the effects on tax revenue are estimated by an average tax rate times the effect on earnings estimated in section VI.A. We use a societal discount rate of 3 percent and a tax distortion rate of 50 percent in the base scenario (see e.g. Heckman et al., 2010, for a similar approach). The time horizon is the 18 years over which outcomes are measured.

The results of this cost-benefit analysis are shown in Table VI. In the base scenario, shown in the first row, the extended language course has a net present value of $40,000 per participant and the benefit-cost ratio is equal to 15. The investment breaks even after 5 years. The next four rows of Table VI show the results when we vary some of the assumptions in the analysis. First, "Alternative price" is based on data from 2008 on prices per language course module instead of prices stated in the law we analyze.[47] Second, we use a discount rate of 7 percent instead of 3 percent in the baseline. Third, we set the tax distortion rate to 0 instead of 50 percent. Finally, we include the cost of additional education obtained by the young cohort of refugees.[48] The net present value is above $22,000, the benefit-cost ratio ranges from 10 to 15 without the cost of formal education but drops to 2 if we include the cost of the additional formal education obtained among the treated. The time to break-even never exceeds 6 years. The reform has similarly large gains for the government due to income tax savings on social welfare payments and added tax payments from the increase in earnings. This also holds when we include the added costs of further education (Appendix Table B.I).

A second finding of the paper is that the temporary (small) reduction in welfare payments to refugees older than 25 in 1999 did not increase their propensity to work nor their labor earnings, but instead may have increased their propensity to commit crimes, especially shoplifting. In this case, while there were some savings from welfare, those should, in principle, be compared with the costs of crime. Since the major part of the increase in crime derives from increases in shoplifting, the costs could be small but are hard to quantify.[49]

The data is more noisy and the sample is smaller when we consider the effects on the second generation. Being able, however, to establish an effect that passes onto the next generation is very powerful as it

[47] Prices were found here on the website of the Danish Ministry of Immigration and Integration: https://integrationsbarometer.dk/danskuddannelserne/takster, checked on June 11 2020.

[48] Section B in the Appendix describes and discusses the assumptions in more detail.

[49] Further details on the cost of crime are found in Section B in the Appendix.

would mean a very long-lasting impact. There are two main reasons why estimated impacts on children are more noisy and uncertain. First, it seems to matter whether one or both parents were treated and, in particular, whether the father received the Danish training too, or only the mother. Second, effects vary by the gender of the child. Boys seem much more sensitive to parents' treatment and possibly especially to treatment of the father in their criminal outcomes. Other researchers focusing on economically disadvantaged students have found that unfavorable family backgrounds are especially harmful for boys (Bertrand and Pan, 2013; Chetty et al., 2016; Autor et al., 2019; Andersen, Dustmann, and Landersø, 2019). Our results broadly confirm those findings, but we acknowledge that the estimates are noisy in this small sample and somewhat sensitive to outliers close to the cutoff. Hence, we should be cautious when interpreting the magnitude of the impact on the second generation. The sign and significance is clear, however. Across estimates, we consistently find that better language skills and labor market integration of the parents improve boys' progress in school and reduce their crime rates. Girls, instead, perhaps because they are at lower risk of performing poorly on these margins, do not seem to show any effect.

Our results suggest that investment in substantial language training of refugees after settlement allows them to be on a better trajectory with higher earnings, higher employment rate and higher propensity to add academic and vocational education later in time. Furthermore, the positive results extend beyond the first generation. There is some evidence that male children of treated parents do better in school and are less likely commit juvenile crimes. From a societal and public finance perspective, this intervention is expensive during the first years but pays off in less than a decade, and produced a substantial return on the investment in the order of fifteen dollars per dollar invested.

References

Agersnap, Ole, Amalie Sofie Jensen, and Henrik Kleven. Forthcoming. "The Welfare Magnet Hypothesis: Evidence From an Immigrant Welfare Scheme in Denmark." *American Economic Review: Insights* .

Andersen, Lars Højsgaard, Christian Dustmann, and Rasmus Landersø. 2019. "Lowering Welfare Benefits: Intended and Unintended Consequences for Migrants and their Families." *CReAM Discussion Paper* (05/19).

Arendt, Jacob Nielsen. 2018. "Integration and Permanent Residence Policies – A Comparative Pilot Study." Study Paper No. 130, The Rockwool Foundation's Research Unit.

———. 2019. "Employment Effects of a Job-First Policy for Refugees." Study Paper No. 139, The Rockwool Foundation's Research Unit.

———. 2020. "The effect of welfare benefit reductions on the integration of refugees." Study Paper No. 151, The Rockwool Foundation's Research Unit.

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. 2019. "Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes." *American Economic Journal: Applied Economics* 11 (3):338–381.

Bansak, Kirk, Jens Hainmueller, and Dominik Hangartner. 2016. "How Economic, Humanitarian, and Religious Concerns Shape European Attitudes Toward Asylum Seekers." *Science* 354 (6309):217–222.

BBC News. 2018. "Reality Check: Are migrants driving crime in Germany?" URL `https://www.bbc.com/news/world-europe-45419466`.

Bertrand, Marianne and Jessica Pan. 2013. "The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior." *American Economic Journal: Applied Economics* 5 (1):32–64.

Bloemraad, Irene and Els De Graauw. 2012. "Immigrant Integration and Policy in the United States: A Loosely Stitched Patchwork." *International Perspectives: Integration and Inclusion* :205–232.

Borjas, George J. 1994. "The Economics of Immigration." *Journal of Economic Literature,* XXXII:1667–1717.

Bratsberg, Bernt, Oddbjørn Raaum, and Knut Røed. 2017. "Immigrant Labor Market Integration Across Admission Classes." *Nordic Economic Policy Review* :17–54.

Brell, Courtney, Christian Dustmann, and Ian Preston. 2020. "The Labor Market Integration of Refugee Migrants in High-Income Countries." *Journal of Economic Perspectives* 34 (1):94–121.

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Roco Titiunik. 2019. "Regression Discontinuity Designs Using Covariates." *The Review of Economics and Statistics* 101 (3):442–451.

Card, David, Jochen Kluve, and Andrea Weber. 2010. "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations." *Journal of the European Economic Association* 16 (3):452–477.

———. 2018. "Active Labor Market Program Evaluations: A Meta-Analysis." *The Economic Journal* 120 (November):894–931.

Chetty, Raj, Nathaniel Hendren, Frina Lin, Jeremy Maherovitz, and Benjamin Scuderi. 2016. "Childhood Environment and Gender Gaps in Adulthood." *American Economic Review: Papers & Proceedings* 106 (5):282–288.

Chiswick, Barry R. 1991. "Speaking, Reading, and Earnings Among Low-Skilled Immigrants." *Journal of Labor Economics* 9:149–170.

Chiswick, Barry R. and P. W. Miller. 1995. "The Endogeneity between Language and Earnings: International Analyses." *Journal of Labor Economics* 13 (3):246–288.

De Vries, Karin. 2013. *Integration at the Border: The Dutch Act on Integration Abroad and International Immigration Law*. Bloomsbury Publishing.

DiPrete, Thomas A and Claudia Buchmann. 2013. *The Rise of Women: The Growing Gender Gap in Education and What It Means for American Schools*. Russell Sage Foundation.

Dustmann, Christian. 1994. "Speaking Fluency, Writing Fluency and Earnings of Migrants." *Journal of Population Economics* 7 (2):133–156.

Fasani, Francesco, Tommaso Frattini, and Luigi Minale. 2018. "(The Struggle for) Refugee Integration into the Labour Market: Evidence from Europe." *IZA Discussion Paper* (11333).

Gelman, Andrew and Guido Imbens. 2019. "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs." *Journal of Business & Economic Statistics* 37 (3):447–456.

Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz. 2010. "The Rate of Return to the HighScope Perry Preschool Program." *Journal of Public Economics* 94 (1-2):114–128.

Hübschmann, Zuzanna. 2015. "Migrant Integration Programs: The Case of Germany." *The Graduate Institute of International and Development Studies, Global Migration Centre* .

Huynh, Duy T., Marie Louise Schultz-Nielsen, and Torben Tranæs. 2007. "The Employment Effects upon Arrival of Reducing Welfare to Refugees." Study Paper No. 15, The Rockwool Foundation's Research Unit.

Hvidtfeldt, Camilla, Marie Louise Schultz-Nielsen, Erdal Tekin, and Mogens Fosgerau. 2018. "An Estimate of the Effect of Waiting Time in the Danish Asylum System on Post-Resettlement Employment Among Refugees: Separating the Pure Delay Effect from the Effects of the Conditions Under which Refugees are Waiting." *PLoS ONE* (13):1–14.

Joona, Pernilla Andersson and Lena Nekby. 2012. "Intensive Coaching of New Immigrants: An Evaluation Based on Random Program Assignment." *The Scandinavian Journal of Economics* 114 (2):575–600.

Joppke, Christian. 2007. "Beyond National Models: Civic Integration Policies for Immigrants in Western Europe." *West European Politics* 30 (1):1–22.

Lee, David S. and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2):281–355.

Lochmann, Alexia, Hillel Rapoport, and Biagio Speciale. 2019. "The Effect of Language Training on Immigrants' Economic Integration: Empirical Evidence from France." 113:165–296. European Economic Review.

LoPalo, Melissa. 2019. "The effects of cash assistance on refugee outcomes." *Journal of Public Economics* 170 (C):27–52.

Martin, Iván, Albert Arcarons, Jutta Aumüller, Pieter Bevelander, Henrik Emilsson, Sona Kalantaryan, Alastair MacIver, Isilda Mara, Giulia Scalettaris, Alessandra Venturini, H Vidovic, I Welle, M Windish, R Wolffberg, and A Zorlu. 2016. "From Refugees to Workers: Mapping Labour Market Integration Support Measures for Asylum-seekers and Refugees in EU Member States." *Volume II: Literature review and country case studies* .

Ministry of Interior Affairs. 1999. “Rapport om resultaterne af Indenrigsministeriets sagsbaserede evaluering af integrationsloven.” .

Rambøll and the Danish Center for Social Science Research. 2013. “Samfundsøkonomisk analyse af metoder. Hjemløsestrategien.” Appendix 3 (Report for the Danish Ministry of Social Affairs and the Interior). URL `https://socialstyrelsen.dk/filer/voksne/hjemloshed/samfundsoekonomisk_analyse.pdf`.

Rosholm, Michael and Rune Vejlin. 2010. “Reducing Income Transfers to Refugee Immigrants: Does Start-Help Help you Start?” *Labour Economics* 17 (1):258–275.

Sarvimäki, Matti and Kari Hämäläinen. 2016. “Integrating Immigrants: The Impact of Restructuring Active Labor Market Programs.” *Journal of Labor Economics* 34 (2):479–508.

Schultz-Nielsen, Marie Louise. 2017. “Labor Market Integration of Refugees in Denmark.” *Nordic Economic Policy Review* :55 – 90.

Williamson, Abigail Fisher. 2018. *Welcoming New Americans? Local Governments and Immigrant Incorporation*. University of Chicago Press.

Tables and Figures

Table I: Summary Statistics of Main Outcomes

	Mean (1)	S.D. (2)
Annual Employment	0.20	0.26
Annual Earnings	8.60	12.41
Obtained Education in Denmark, Year 18	0.13	0.33
Communicative Relative to Manual Tasks	-1.07	0.90
Top-Three Occupations:		
Personal and Protective Services	0.31	0.46
Sales and Services	0.31	0.46
Machine Operating and Assembling	0.21	0.40
Not Living in Initial Municipality, Year 18	0.57	0.49
Living in Urban Municipality, Year 18	0.55	0.50
Annual Convictions of Crime	0.03	0.07
Annual Convictions of Property Crime	0.02	0.06
Annual Convictions of Shoplifting (Supermarket)	0.01	0.03

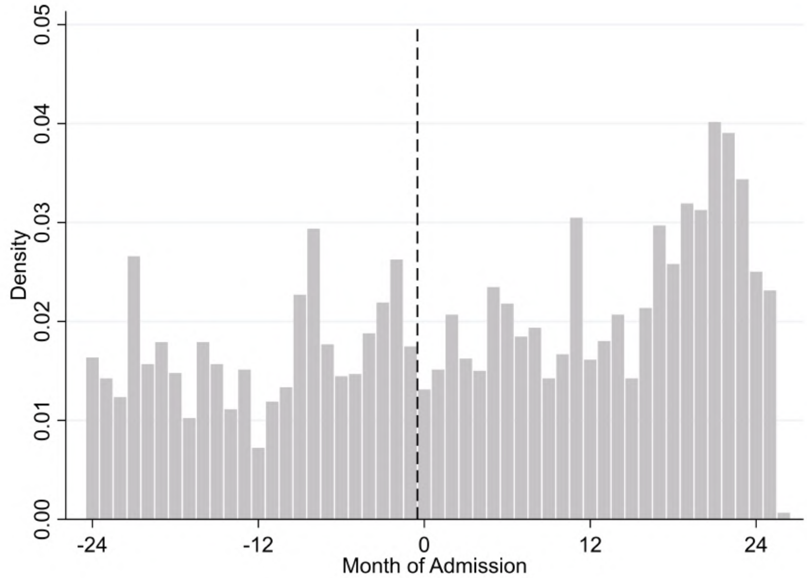
Notes: Earnings are expressed in thousands of US Dollars (at year 2000 prices). Employment is measured as a fraction of full-time working year. The ratio of task intensities is calculated as log(communicative tasks)-log(manual tasks). The variables relative to "obtained education" and location of living are measured in year 2016, 18 years after the reform. The remaining variables are sample means for the 18 year period from Jan. 1999 to Dec.2016. Our sample is constituted by 8,558 refugees. Task intensity and occupations are conditional on being employed with a valid ISCO code (N=5,195 and N=5,122 can be linked to task data).

Table II: Language Training and Employment Support the First Three Years After Arrival

	1 Year (1)	1.5 Years (2)	2 Years (3)	2.5 Years (4)	3 Years (5)
	Panel a. Share Ever Participating in Language Training				
RD estimate	0.053	0.041	-0.025	0.001	0.011
	(0.064)	(0.048)	(0.037)	(0.032)	(0.030)
Mean of Untreated at Cutoff	0.806	0.854	0.915	0.906	0.919
	Panel b. Share Ever Participating in Employment Support				
RD estimate	0.160*	0.030	-0.038	-0.066	-0.081
	(0.085)	(0.069)	(0.065)	(0.063)	(0.064)
Mean of Untreated at Cutoff	0.280	0.423	0.529	0.596	0.654
N	6,868	7,631	8,558	8,558	8,558
	Panel c. Cumulative Hours of Language Training				
RD estimate	104.987**	147.946**	228.432***	253.201***	186.853*
	(50.846)	(60.937)	(75.357)	(94.297)	(97.930)
Mean of Untreated at Cutoff	355.575	517.920	662.987	789.475	981.739
	Panel d. Cumulative Hours of Employment Support				
RD estimate	73.475	-129.844	-61.183	-39.840	-36.505
	(159.041)	(98.478)	(101.132)	(136.952)	(148.881)
Mean of Untreated at Cutoff	512.459	811.708	948.167	1156.387	1347.726
N	4,446	5,028	5,576	5,996	6,324

Notes: $*p = 0.10, **p = 0.05, ***p = 0.01$. Robust standard errors in parenthesis. Each panel reports the RD estimate and the mean of the variable for untreated measured at the cutoff corresponding to α from equation (1). Estimates are obtained from local linear regressions using the triangular kernel and the optimal bandwidth selector from Calonico et al. (2019).

(a) Distribution of Admissions



(b) Demeaned Monthly Share of Annual Admissions

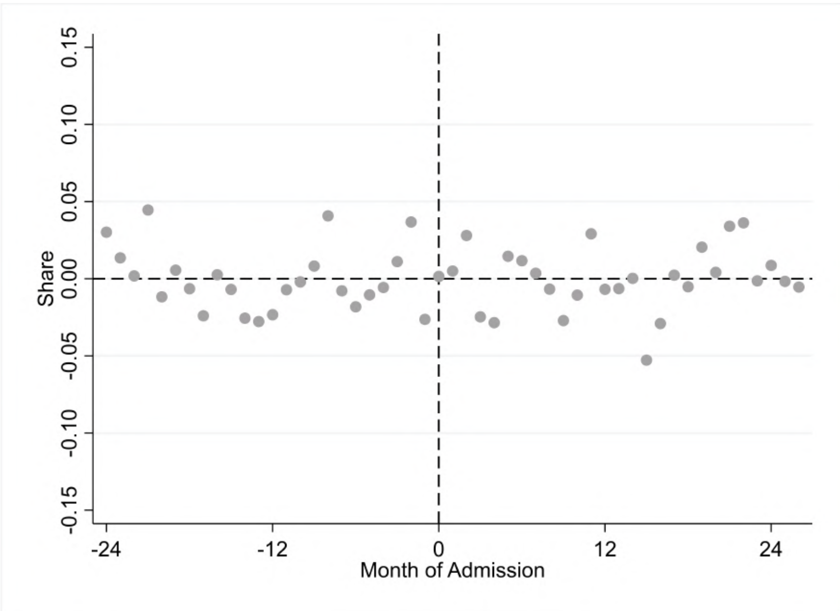


Figure I: Refugees by Month of Receiving Residency

Notes: Admissions from January 1997 to February 2001. Panel a shows the density and Panel b shows the share of refugees in each month demeaned by the average monthly share of annual admissions, excluding from the mean the 6 months before and after the reform.

Table III: Summary Statistics and Balancing Tests

	Mean (1)	S.D. (2)	RD Estimate (3)	Confidence Interval (4)
Panel a. Individual Characteristics				
Age	31.50	7.68	-0.39	[-1.71 ; 0.92]
Married	0.67	0.47	0.05	[-0.04 ; 0.13]
Female	0.42	0.49	0.04	[-0.04 ; 0.13]
No. Children $<$ 3y	0.19	0.43	0.03	[-0.05 ; 0.10]
No. Children 3-17y	1.00	1.51	-0.04	[-0.42 ; 0.34]
Iraq	0.44	0.50	-0.03	[-0.14 ; 0.08]
Afghanistan	0.16	0.36	-0.02	[-0.11 ; 0.07]
Other Country	0.40	0.49	0.03	[-0.10 ; 0.15]
Speaks Arabic	0.35	0.48	-0.04	[-0.13 ; 0.05]
Speaks Dari	0.09	0.29	-0.01	[-0.05 ; 0.03]
Danish 1	0.22	0.42	-0.07	[-0.19 ; 0.05]
Danish 2	0.35	0.48	-0.00	[-0.09 ; 0.08]
Danish 3	0.32	0.46	0.05	[-0.04 ; 0.15]
Quota Refugee	0.10	0.30	-0.03	[-0.08 ; 0.02]
Convention Refugee	0.18	0.38	-0.08	[-0.19 ; 0.02]
Family-Reunified	0.19	0.40	0.03	[-0.07 ; 0.13]
Other Refugee	0.53	0.50	0.01	[-0.13 ; 0.14]
Wait Time Asylum (Days)	404.33	348.11	16.58	[-72.41 ; 105.56]
Panel b. Municipality Characteristics				
Urban Municipality	0.33	0.47	0.11	[-0.01 ; 0.24]
Five Largest Cities	0.17	0.37	0.01	[-0.08 ; 0.10]
Employment Rate 1996	0.74	0.04	-0.00	[-0.01 ; 0.01]
Unemployment Rate 1996	0.09	0.02	0.00	[-0.00 ; 0.01]
Avg. Labor Income 1996	28.52	3.69	0.43	[-0.37 ; 1.23]
Number of Co-Nationals	222.09	679.63	-23.55	[-180.01 ; 132.91]
Share of Co-Nationals	0.00	0.00	0.00	[-0.00 ; 0.00]

Notes: Summary statistics (columns 1-2) and balancing tests (columns 3-4) of the impact of the reform on predetermined variables for refugees obtaining refugee status in Denmark between January 1997 and December 2000. The RD estimates are from local linear regressions using the triangular kernel and the optimal bandwidth selector from Calonico et al. (2019). The confidence intervals are constructed based on robust standard errors. Age, marital status and the number of children are measured at date of immigration. Danish 1 to 3 refer to the language track the individual was initially placed in. Wait time for asylum is the number of days between application and admission, and it is calculated for refugees (excluding quota refugees). Municipality characteristics refer to the municipality the refugee is placed in after admission. Urban municipalities are municipalities in the capital area or municipalities with a town of more than 45,000 inhabitants. The five largest cities are Copenhagen (including Frederiksberg Municipality), Aarhus, Odense, Aalborg and Esbjerg. Average income in the municipality is measured in 1,000 USD (2000 level). The number of observations is 8,558, except for waiting time which is calculated for 5,956 refugees (not applicable for family-reunification or UNHCR quota refugees).

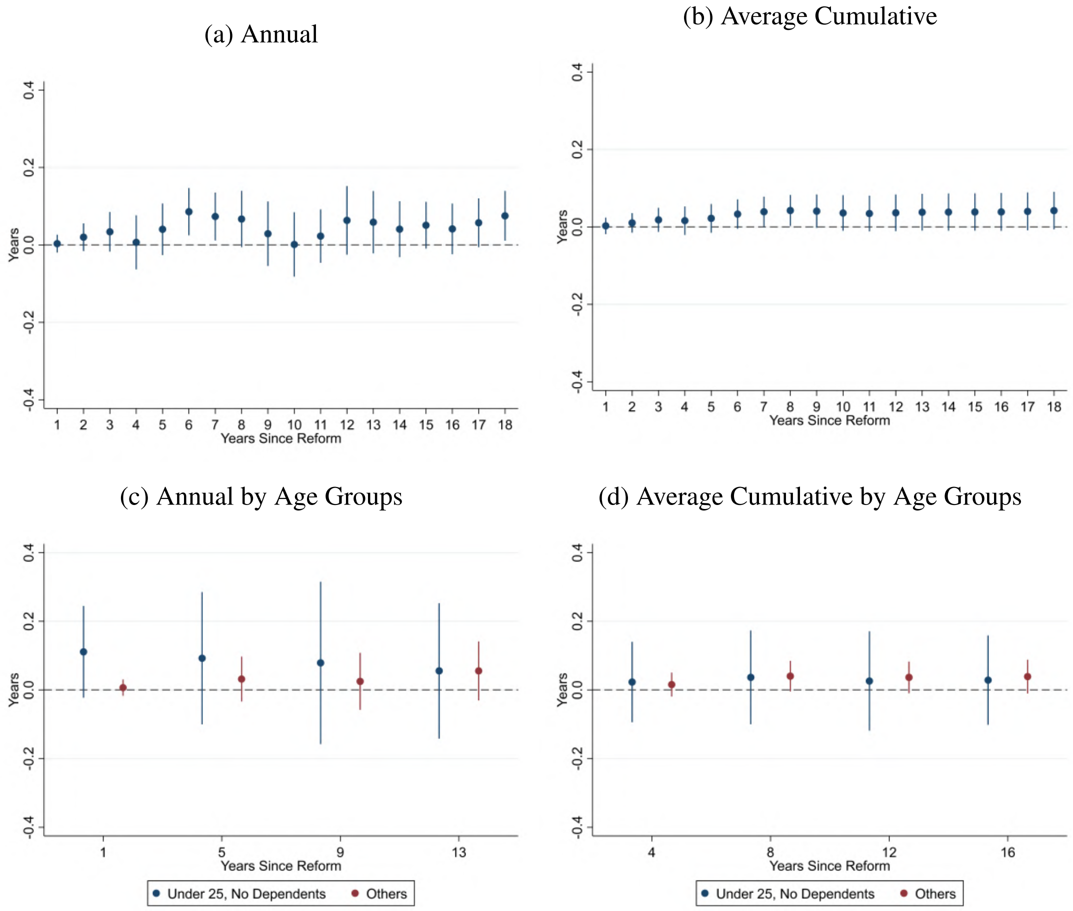
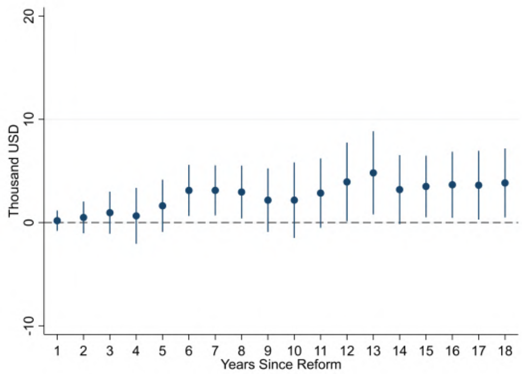


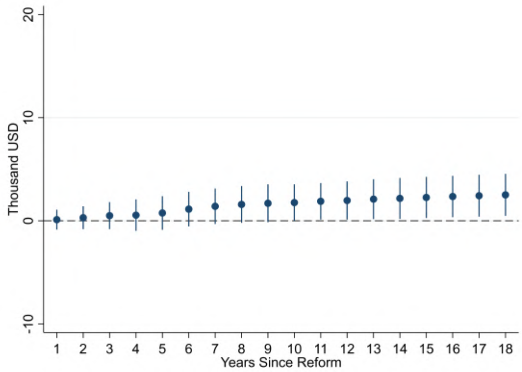
Figure II: Employment Effects

Notes: RD estimates and 95-percent confidence intervals based on robust standard errors from local linear regressions using the triangular kernel and the optimal bandwidth selector from Calonico et al. (2019). Panels a and c show annual employment, and Panels b and d show cumulative employment divided by the number of years since the reform. In Panels c and "others" include persons under 25 with children and all individuals above 25. The employment rate is measured as a fraction of a full-time working year. This measure takes the value one if the worker was a full time employee during the whole year while it is less than one if the individual was either a part-time employee or not employed at some point throughout the year.

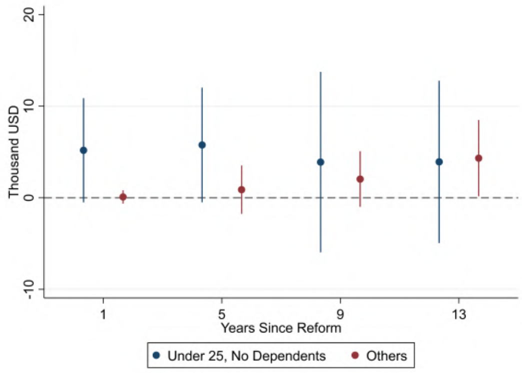
(a) Annual



(b) Average Cumulative



(c) Annual by Age Groups



(d) Average Cumulative by Age Groups

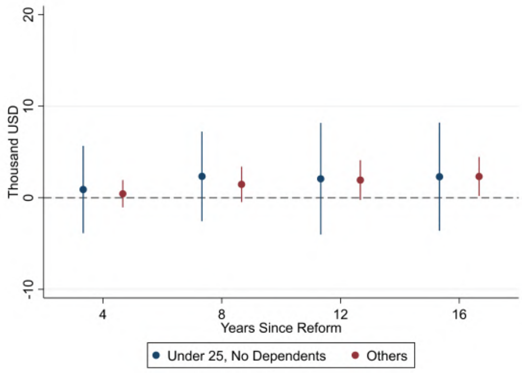
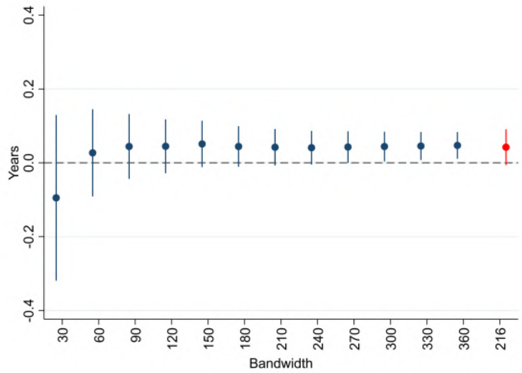


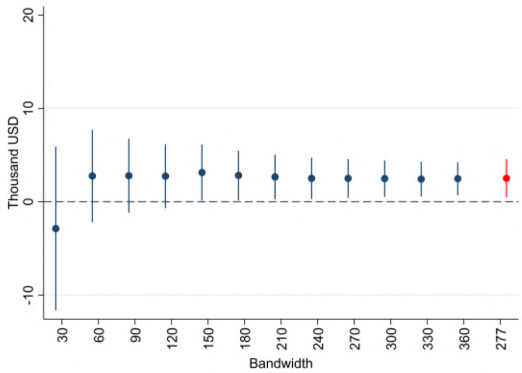
Figure III: Earnings Effects

Notes: RD estimates and 95-percent confidence intervals based on robust standard errors from local linear regressions using the triangular kernel and the optimal bandwidth selector from Calonico et al. (2019). Panels a and c show annual earnings, and Panels b and d show cumulative earnings divided by the number of years since the reform. In Panels c and d “others” include persons under 25 with children and all individuals above 25.

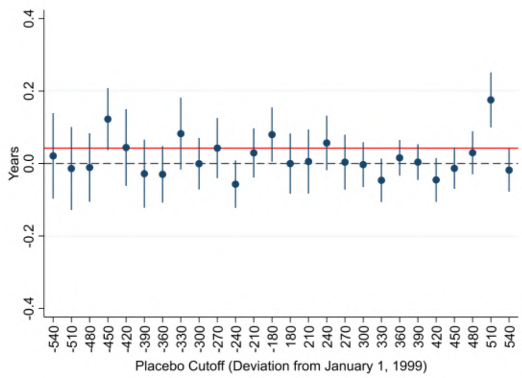
(a) Bandwidth, Average Cumulative Employment



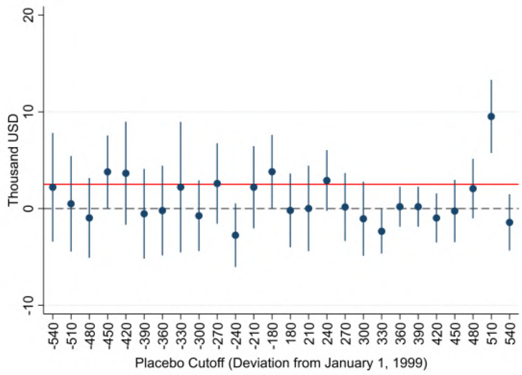
(b) Bandwidth, Average Cumulative Earnings



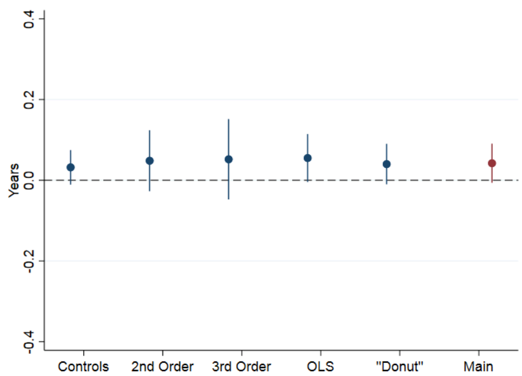
(c) Placebo Cutoffs, Average Cumulative Employment



(d) Placebo Cutoffs, Average Cumulative Earnings



(e) Functional Form, Average Cumulative Employment



(f) Functional Form, Average Cumulative Earnings

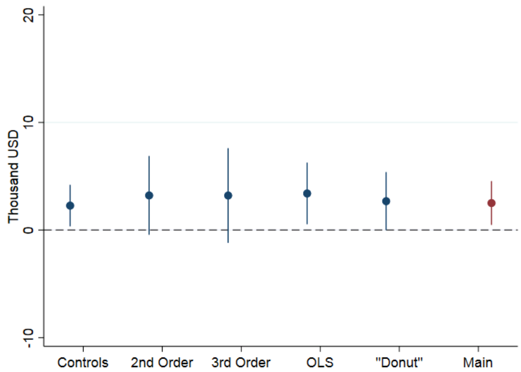
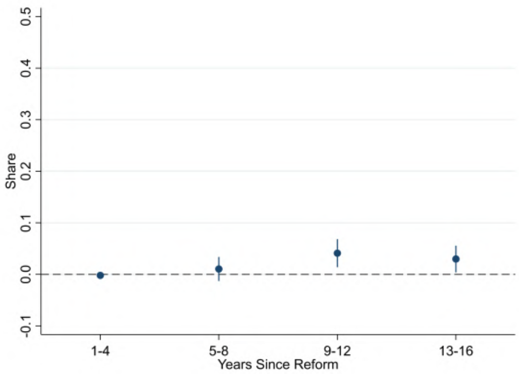


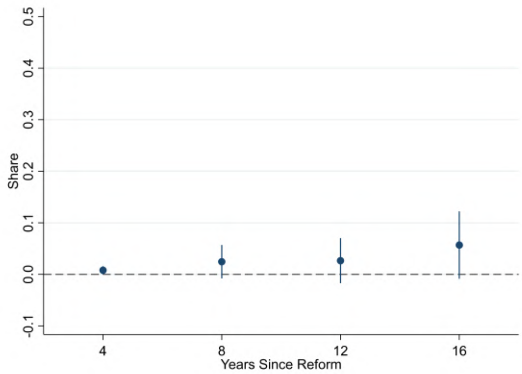
Figure IV: Robustness Checks, Average Cumulative Employment and Earnings

Notes: Panels a and b show the sensitivity to the choice of bandwidth. The red dot and bar are the estimate and 95-percent confidence interval based on robust standard errors from year 18 in Panel b of Figures II and III using the optimal bandwidth from Calonico et al. (2019) Panel c and d examine estimates at made-up cutoff points (red horizontal line is the estimate from Panel b of Figures II and III). Panels e and f compare the main specification to estimates including control variables, using 2nd or 3rd order polynomials or OLS, and excluding refugees admitted in the four weeks immediately around the cutoff ("donut"). Control variables are age, age squared, unmarried, female, number of children between 0-2 years old and 3-17 years old, Iraq, Afghanistan, speaks Arabic, speaks Dari, Danish 1, 2 or 3 (unknown level is the reference), quota refugee, family-reunified or other refugee (convention refugee is the reference).

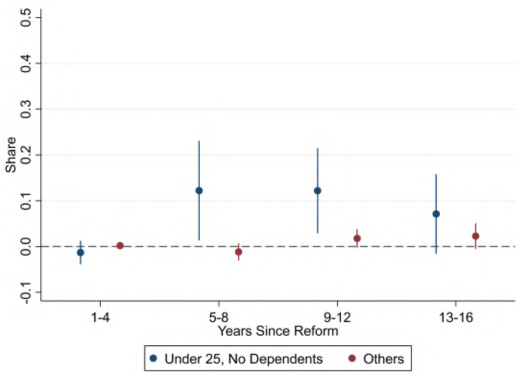
(a) Enrolled in Education in Denmark



(b) Obtained Education in Denmark



(c) Enrolled in Education in Denmark by Age Groups



(d) Obtained Education in Denmark by Age Groups

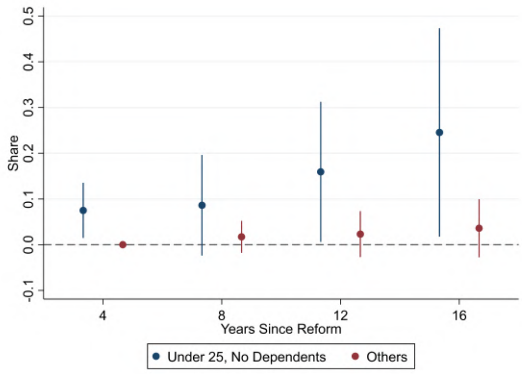
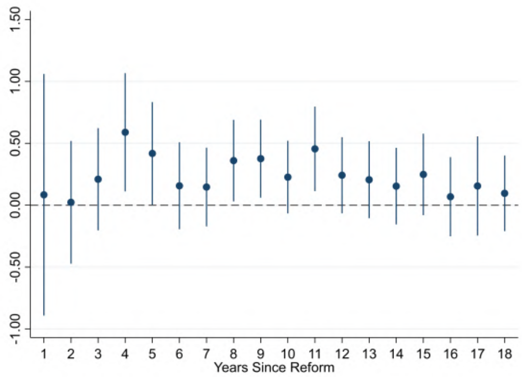


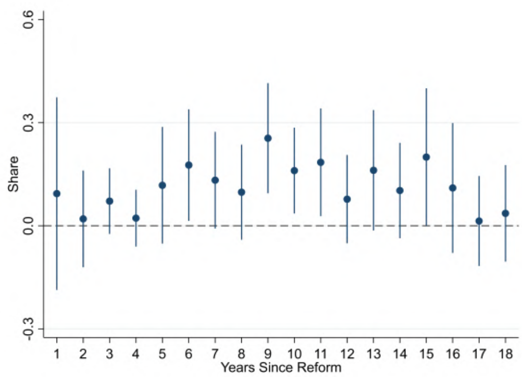
Figure V: Educational Attainment in Denmark

Notes: RD estimates and 95-percent confidence intervals based on robust standard errors from local linear regressions using the triangular kernel and the optimal bandwidth selector from Calonico et al. (2019). Panels a and b show, respectively, the impact on enrollment in and obtaining education in Denmark. Panels c and d show the same outcomes by age groups. “Others” include persons under 25 with children and all individuals above 25.

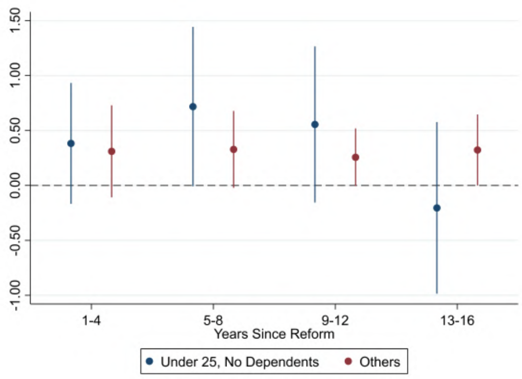
(a) Communicative Relative to Manual Tasks



(b) Personal and Protective Services



(c) Communicative Relative to Manual Tasks by Age Groups



(d) Personal and Protective Services by Age Groups

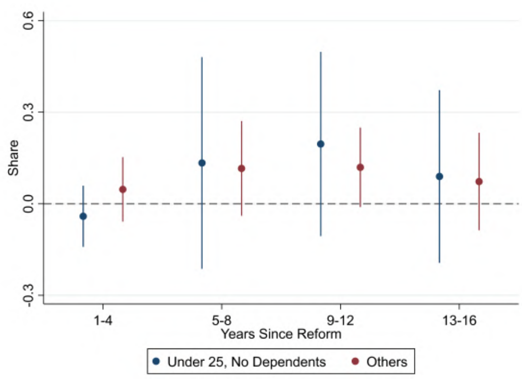
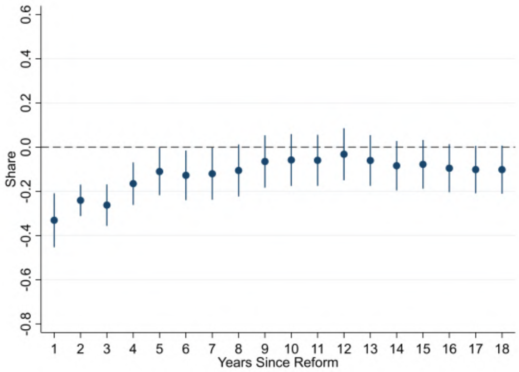


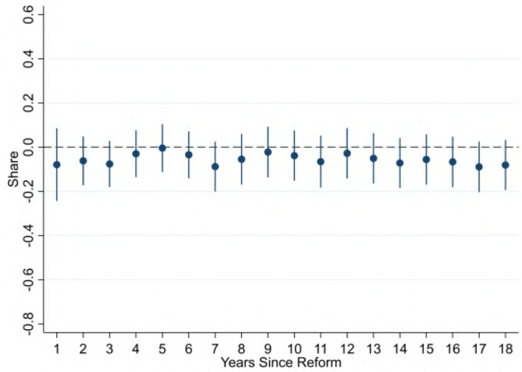
Figure VI: Specialization in the Labor Market

Notes: RD estimates and 95-percent confidence intervals based on robust standard errors from local linear regressions using the triangular kernel and the optimal bandwidth selector from Calonico et al. (2019). Panels a and b show, respectively, the impact on the log(communicative tasks)-log(manual tasks) and working in personal and protective services. Panels c and d show the same outcomes by age groups. “Others” include persons under 25 with children and all individuals above 25.

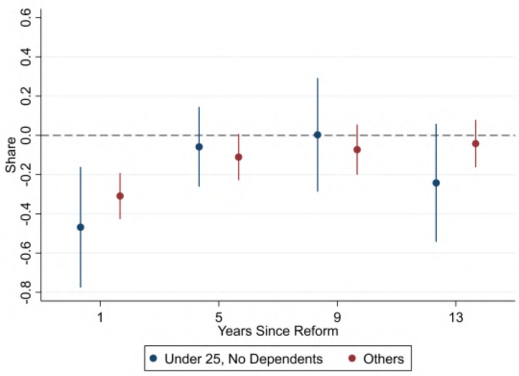
(a) Not Living in Initial Municipality



(b) Living in Urban Municipality



(c) Not Living in Initial Municipality by Age Groups



(d) Living in Urban Municipality by Age Groups

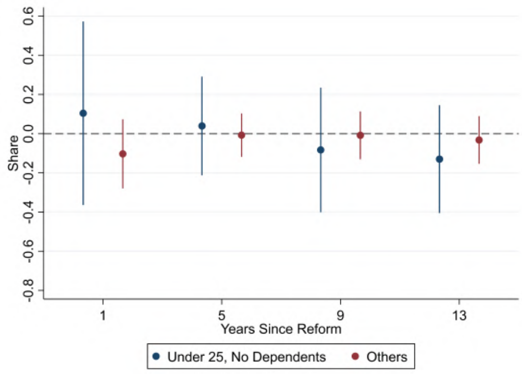
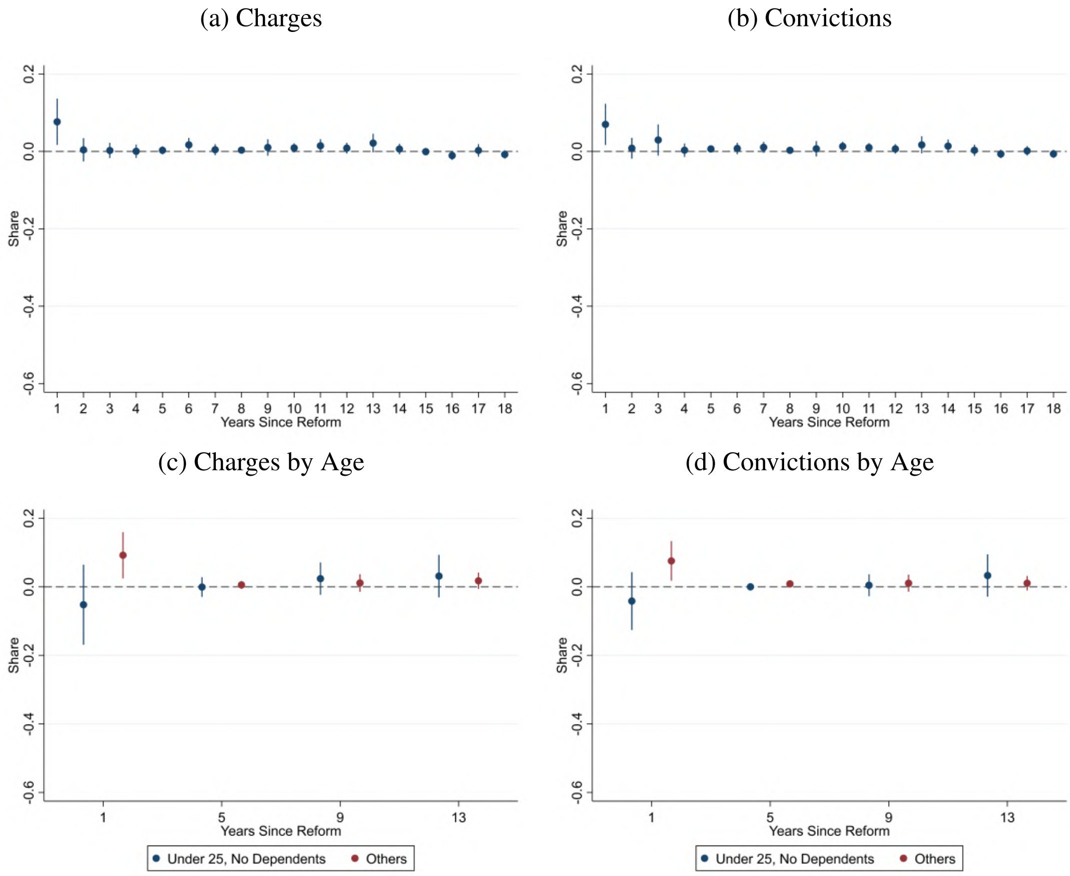


Figure VII: Mobility Effects

Notes: RD estimates and 95-percent confidence intervals based on robust standard errors from local linear regressions using the triangular kernel and the optimal bandwidth selector from Calonico et al. (2019). Panels a and b show the share not living in their initial municipality and the share living in an urban municipality. Panels c and d show the same outcomes by age groups. “Others” include persons under 25 with children and all individuals above 25.



(a) Charges

(b) Convictions

(c) Charges by Age

(d) Convictions by Age

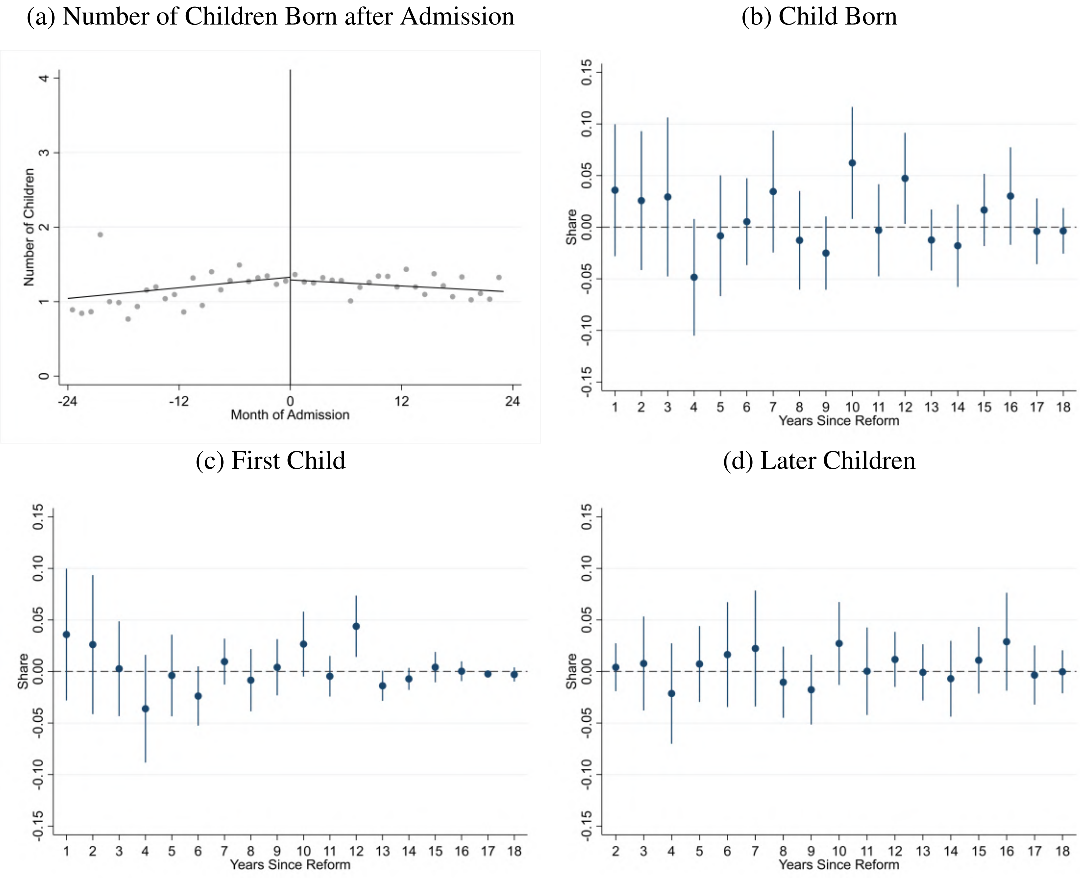
Figure VIII: Shoplifting in Supermarkets

Notes: RD estimates and 95-percent confidence intervals based on robust standard errors from local linear regressions using the triangular kernel and the optimal bandwidth selector from Calonico et al. (2019). In Panels c and d "others" include persons under 25 with children and all individuals above 25.

Table IV: Heterogeneous Effects, Sub-Sample Analysis

	Average Cumulative Employment (1)	Average Cumulative Earnings (2)	Obtained Education (3)	Charged Shoplifting (4)	Convicted Shoplifting (5)
			Panel a. Gender		
Female	0.058*	2.978**	0.091	0.104**	0.075*
	(0.034)	(1.244)	(0.061)	(0.050)	(0.042)
Male	0.035	2.534	-0.009	0.042	0.061**
	(0.036)	(1.663)	(0.038)	(0.029)	(0.030)
			Panel b. Alphabet of Mother Tongue		
Arabic Alphabet	0.077**	3.810**	0.059	0.036	0.043
	(0.034)	(1.516)	(0.057)	(0.032)	(0.033)
Other Alphabet Than Arabic	0.006	0.821	0.056	0.102**	0.089**
	(0.035)	(1.501)	(0.046)	(0.040)	(0.036)

Notes: $*p = 0.10, **p = 0.05, ***p = 0.01$. Robust standard errors in parenthesis. Each column refers to a different outcome variable. Columns 1-3 are measured in year 18 and columns 4-5 in year 1. Each row shows RD estimates from a regression on a sub-sample using the optimal bandwidth from the full sample.



(a) Number of Children Born after Admission (b) Child Born

(c) First Child (d) Later Children

Figure IX: Fertility

Notes: Panel a shows the total number of children born after admission. Panels b to d show post admission fertility RD estimates and 95-percent confidence intervals based on robust standard errors from local linear regressions using the triangular kernel and the optimal bandwidth selector from Calonico et al. (2019).

Table V: Education and Criminal Outcomes of Children

	Boys				Girls			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Took Any Exam in Lower Secondary School	0.163**	0.137*	0.238***	0.162***	0.073	0.024	-0.020	-0.006
	(0.071)	(0.082)	(0.059)	(0.056)	(0.045)	(0.049)	(0.055)	(0.050)
Graduated Lower Secondary School	0.113**	0.094**	0.103*	0.178***	0.008	-0.037	0.010	-0.009
	(0.055)	(0.047)	(0.059)	(0.066)	(0.075)	(0.071)	(0.061)	(0.063)
Enrolled in Upper Secondary Education	0.066	0.015	0.068	0.089*	-0.027	-0.015	-0.062	-0.081
	(0.053)	(0.054)	(0.051)	(0.053)	(0.073)	(0.072)	(0.060)	(0.064)
Charged with a Crime	-0.159*	-0.236**	-0.061	-0.025	0.001	-0.046	-0.013	0.023
	(0.096)	(0.101)	(0.079)	(0.074)	(0.050)	(0.039)	(0.039)	(0.044)
Convicted of a Crime	-0.111	-0.163	-0.049	-0.003	-0.028	-0.054	-0.025	-0.016
	(0.093)	(0.102)	(0.071)	(0.067)	(0.042)	(0.037)	(0.034)	(0.036)
N	5,542	4,834	5,542	5,140	4,188	3,698	4,188	3,865

Notes: $*p = 0.10, **p = 0.05, ***p = 0.01$. Robust standard errors in parenthesis. RD estimates from local linear regressions using the triangular kernel and the optimal bandwidth selector from Calonico et al. (2019). Children of refugees born before 2003. (1 and 5) uses the admission date of the first parent. (2 and 6) uses the admission date of the father. (3 and 7) uses the admission date of the last parent. (4 and 8) uses the admission date of the mother. The outcomes measuring taking any exam in lower secondary school or graduating are dummies for taking the exam or graduating before age 17. Enrollment in upper secondary school is an indicator for enrolling prior to 2018. The crime outcomes are dummies for ever being charged with or convicted of a crime before 2018.

Table VI: Societal Cost-Benefit Analysis

	B (1)	C (2)	DWL (3)	NPV (4)	$(B + DWL)/C$ (5)	Years Before $NPV > 0$ (6)
Base scenario	35.2	-2.8	7.7	40.1	15.4	5
Alternative price	35.2	-3.9	7.2	38.4	10.8	6
Discount rate 7%	23.0	-2.8	5.0	25.3	10.1	5
Tax distortion rate	35.2	-2.8	0.0	32.4	12.6	5
With cost of education	35.2	-16.0	3.1	22.3	2.4	6

Notes: Columns 1 to 5 are measured in 1,000 USD and column 6 is denoted in years. B is benefits, C is costs, DWL if the deadweight loss (or gain in our setting) from (lower) taxation, and NPV is the net present value of the investment over 18 years.

A Appendix: Additional Tables and Figures

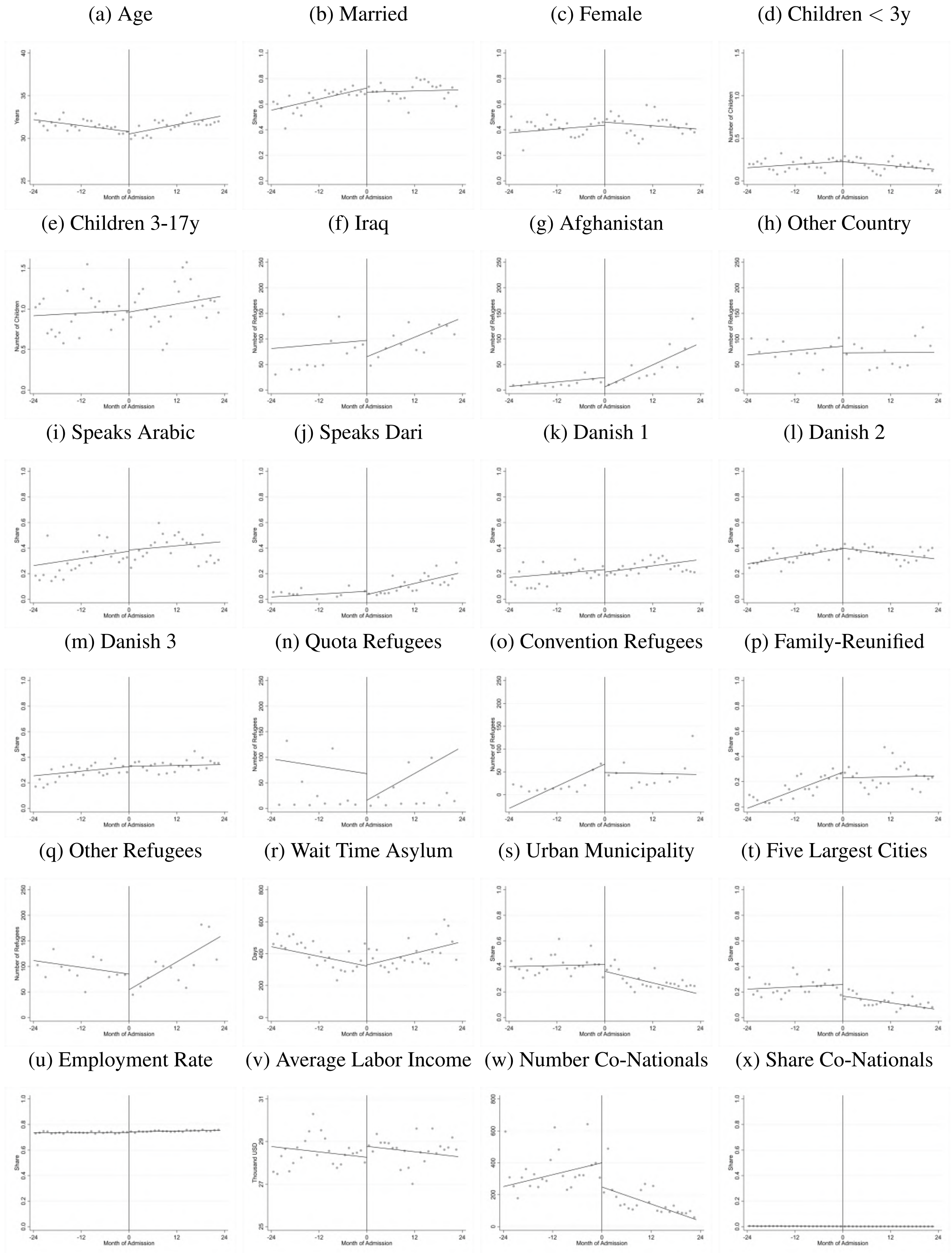


Figure A.I: Predetermined Covariates by Month of Admission

Notes: Each panel shows a predetermined covariate from Table III (we exclude the unemployment rate to fit the figure on one page). The dots correspond to sample means by one-month bins. Bins with less than 5 observations are excluded.

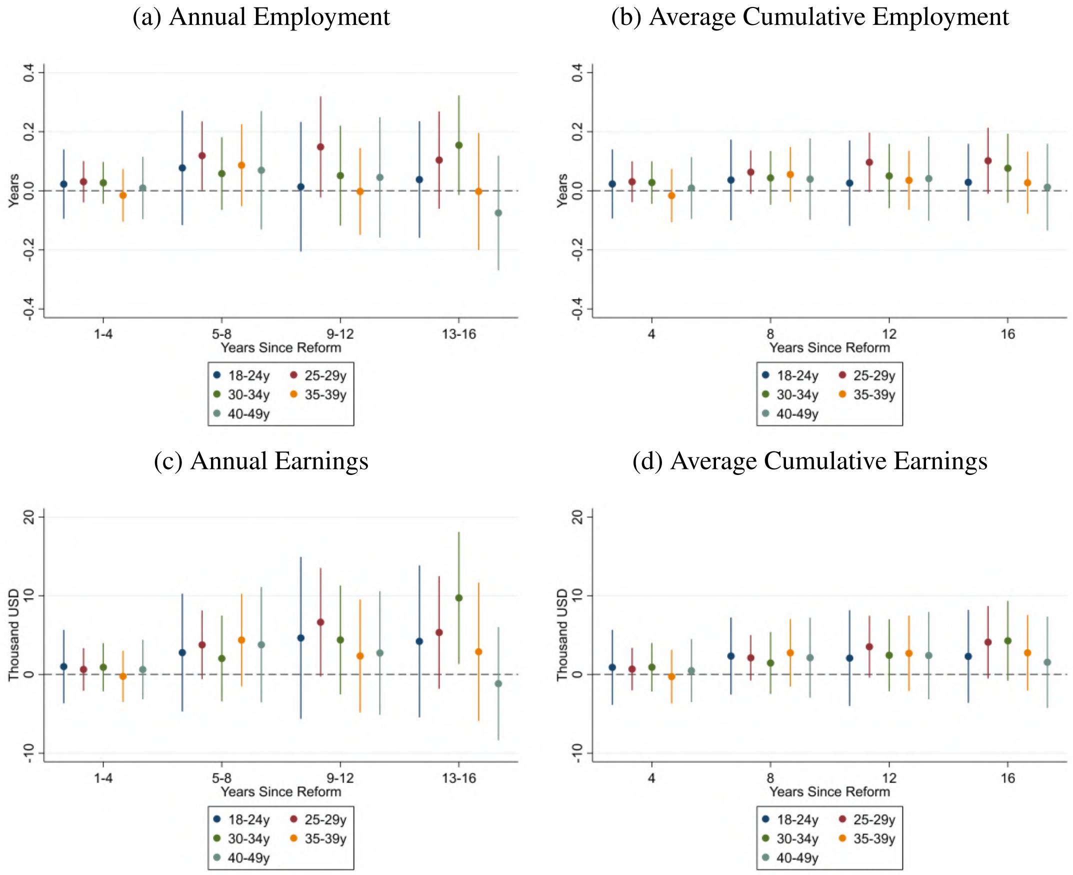


Figure A.II: Employment and Earnings Effects by Age Groups

Notes: RD estimates and 95-percent confidence intervals based on robust standard errors from local linear regressions using the triangular kernel and the optimal bandwidth selector from Calonico et al. (2019). Estimates are shown for refugees without dependents for the 18-24 year age group.

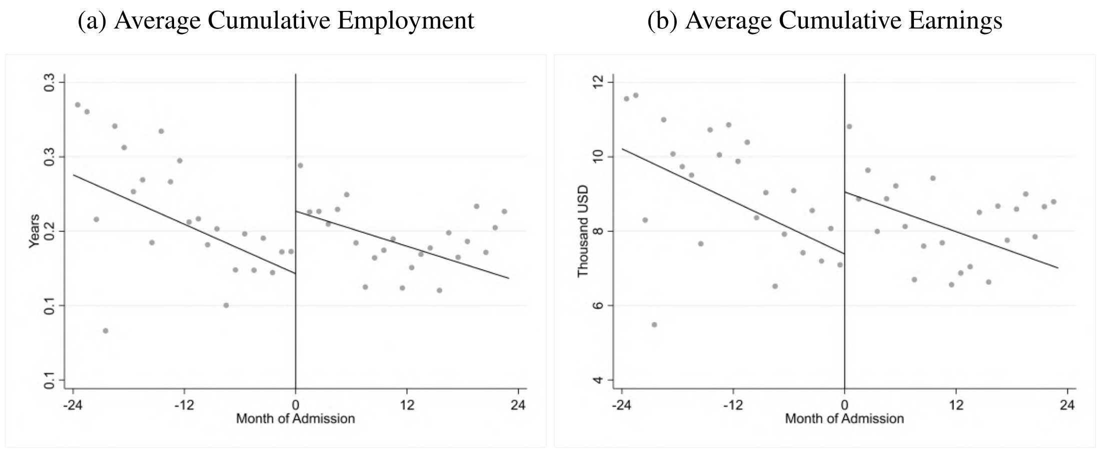


Figure A.III: Average Cumulative Employment and Earnings in Year 18

Notes: The graphs show sample means by one-month bins of average cumulative employment and average cumulative earnings in year 18. Bins with less than 5 observations are excluded.

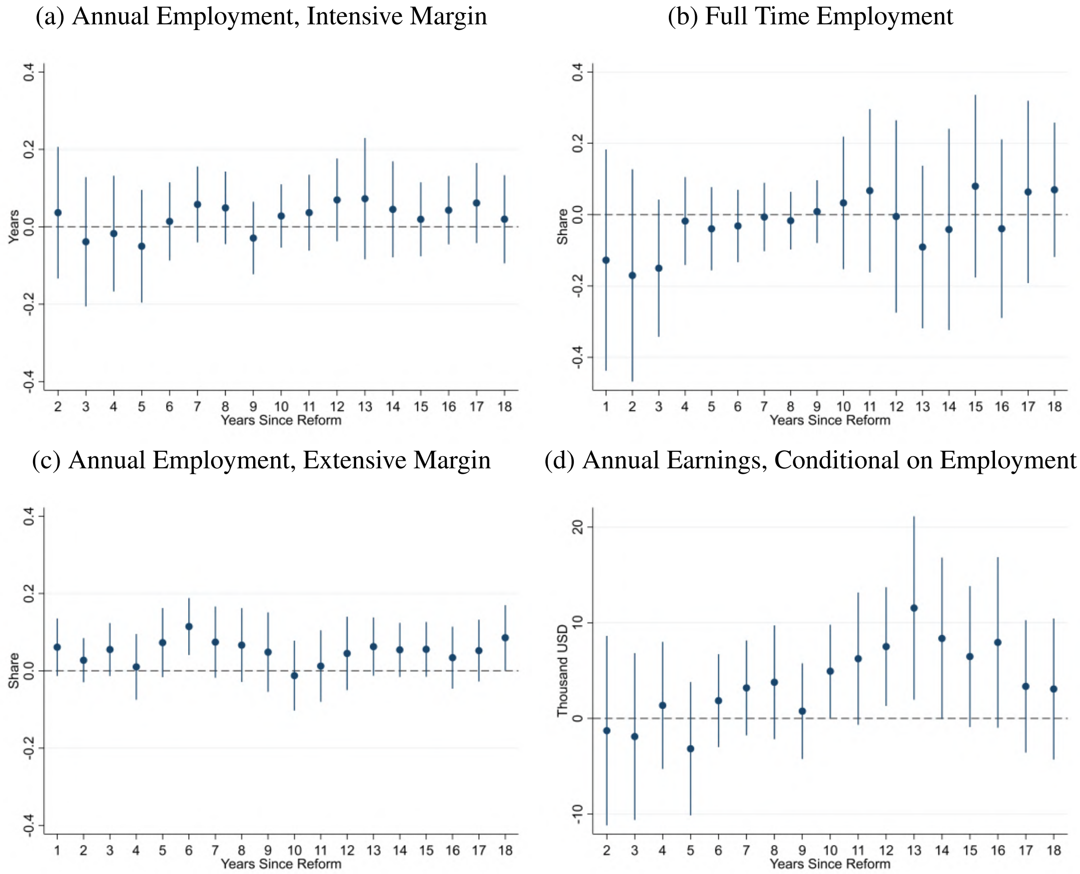
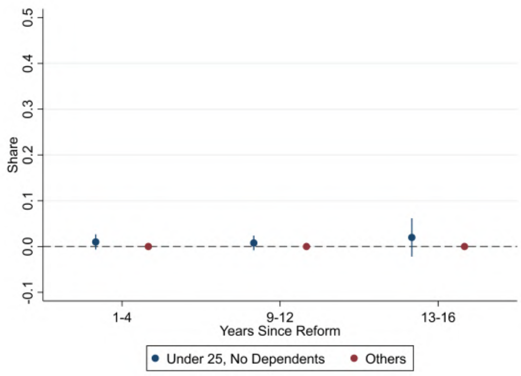


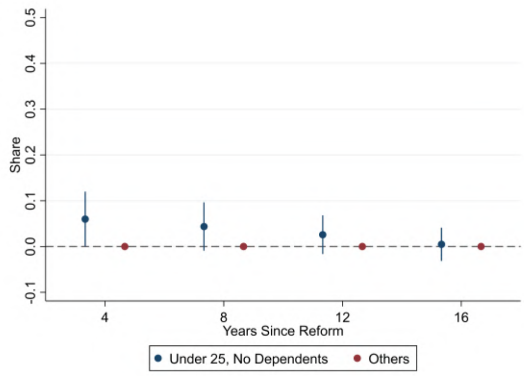
Figure A.IV: Margins of Employment on Earnings Responses

Notes: RD estimates and 95-percent confidence intervals based on robust standard errors from local linear regressions using the triangular kernel and the optimal bandwidth selector from Calonico et al. (2019). Panel a shows the intensive margin response (fraction of a full time year worked conditional on working). Panel b shows the impact on full time employment relative to less than full time employment, conditional on working. Panel c shows the extensive margin employment response and Panel d shows the impacts on annual earnings for those who work.

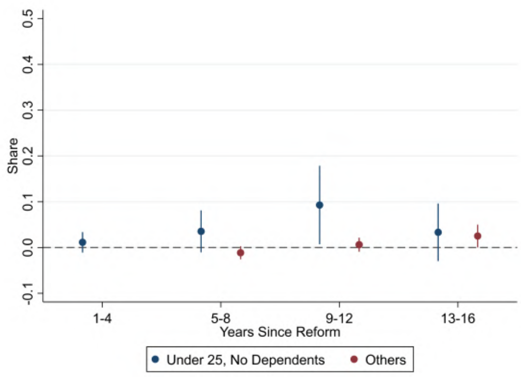
(a) Enrolled in Lower Secondary School



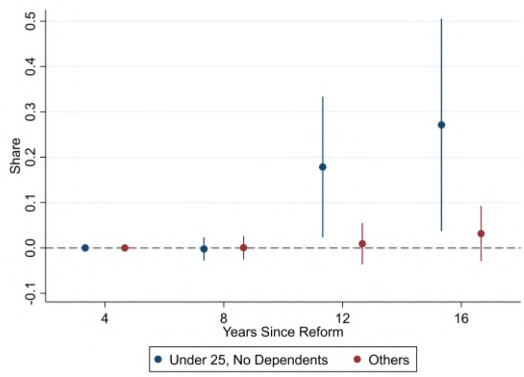
(b) Completed Lower Secondary School



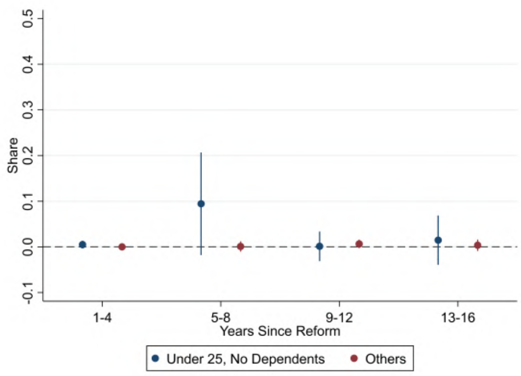
(c) Enrolled in Vocational Education



(d) Completed Vocational Education



(e) Enrolled in Academic Education



(f) Completed Academic Education

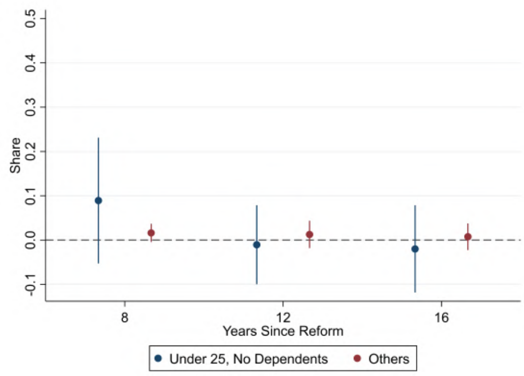


Figure A.V: Level of Educational Attainment in Denmark by Age Groups

Notes: RD estimates and 95-percent confidence intervals based on robust standard errors from local linear regressions using the triangular kernel and the optimal bandwidth selector from Calonico et al. (2019). The results are shown for refugees younger than 25 years without dependents at immigration and others. “Others” include refugees younger than 25 years with dependents at immigration.

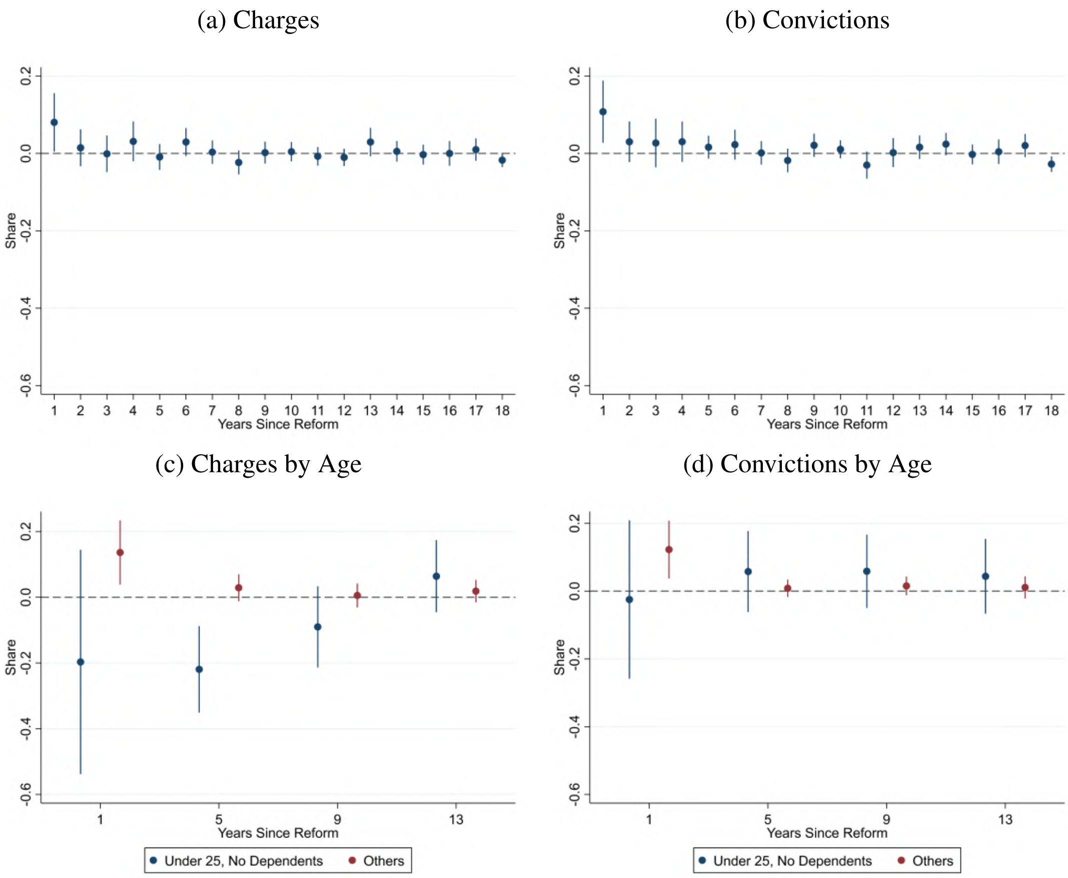


Figure A.VI: All Crimes

Notes: RD estimates and 95-percent confidence intervals based on robust standard errors from local linear regressions using the triangular kernel and the optimal bandwidth selector from Calonico et al. (2019). In Panels c and d "others" include persons under 25 with children and all individuals above 25.

Table A.I: Heterogeneous Effects, Interactions

	Average Cumulative Employment (1)	Average Cumulative Earnings (2)	Obtained Education (3)	Charged Shoplifting (4)	Convicted Shoplifting (5)
			Panel a. Gender		
RD	0.035	2.534	-0.009	0.042	0.061**
	(0.037)	(1.652)	(0.037)	(0.029)	(0.030)
Female	-0.085***	-4.819***	0.090*	0.006	0.024
	(0.032)	(1.191)	(0.050)	(0.024)	(0.024)
Interaction Term	0.024	0.444	0.100	0.062	0.014
	(0.050)	(2.082)	(0.073)	(0.059)	(0.053)
Constant	0.224***	9.781***	0.058**	-0.005	-0.022
	(0.022)	(0.945)	(0.025)	(0.016)	(0.016)
			Panel b. Alphabet of Mother Tongue		
RD	0.006	0.821	0.056	0.102**	0.089**
	(0.036)	(1.509)	(0.047)	(0.042)	(0.036)
Arabic Alphabet	-0.102***	-4.355***	0.053	0.042*	0.047**
	(0.031)	(1.189)	(0.049)	(0.023)	(0.024)
Interaction Term	0.070	2.989	0.003	-0.065	-0.047
	(0.049)	(2.107)	(0.076)	(0.055)	(0.051)
Constant	0.225***	9.308***	0.069**	-0.020	-0.030*
	(0.023)	(0.901)	(0.034)	(0.015)	(0.016)

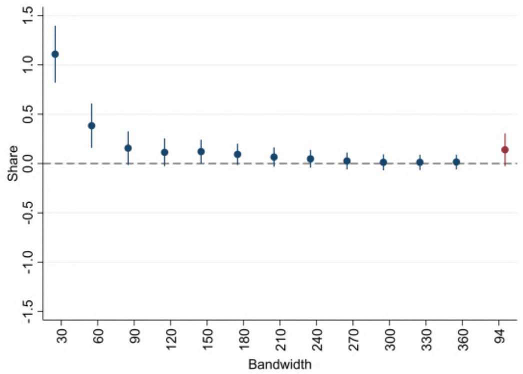
Notes: RD estimates from local linear regressions using the triangular kernel and the optimal bandwidths from the full sample without interaction terms. Robust standard errors in parenthesis ($*p = 0.10, **p = 0.05, ***p = 0.01$). Each column refers to a different outcome variable. Columns 1-3 are measured in year 18 and columns 4-5 in year 1.

Table A.II: Summary Statistics for Children of Refugees

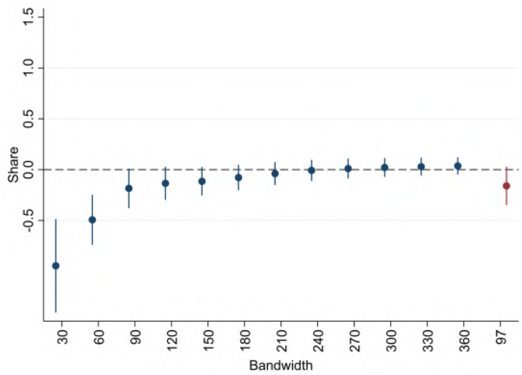
	Panel a. Born After Admission			Panel b. Born Before Admission		
	Mean	S.D.	N	Mean	S.D.	N
Male	0.52	0.50	2,294	0.59	0.49	7,436
Year of Birth	2000	1.33	2,294	1992	3.69	7,436
Age Mother at Childbirth	28.45	5.30	2,013	25.48	5.37	6,992
Age Father at Childbirth	33.12	5.55	1,991	29.93	5.72	6,535
All Parents Refugees or Family-Reunified	0.85	0.36	2,294	0.96	0.19	7,436
Parents Admitted on Same Day	0.20	0.40	2,294	0.25	0.43	7,436
Parents Admitted Before Reform	0.39	0.49	2,294	0.64	0.48	7,436
One Parent Admitted Before Reform	0.17	0.38	2,294	0.07	0.25	7,436
Parents Admitted After Reform	0.44	0.50	2,294	0.29	0.45	7,436

Notes: Panel a shows summary statistics for children born between 1997-2002 after their last refugee parent was admitted. Panel b shows the same statistics for children of refugees born before their parents were admitted.

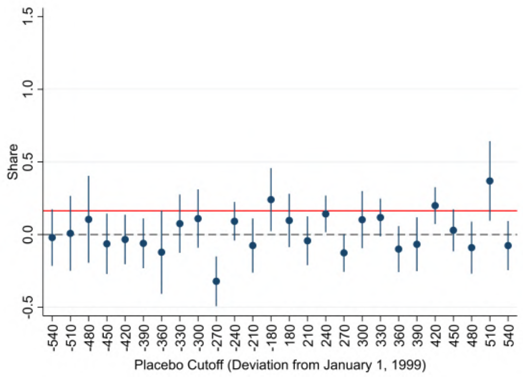
(a) Bandwidth, Took Any Exam in Lower Secondary School



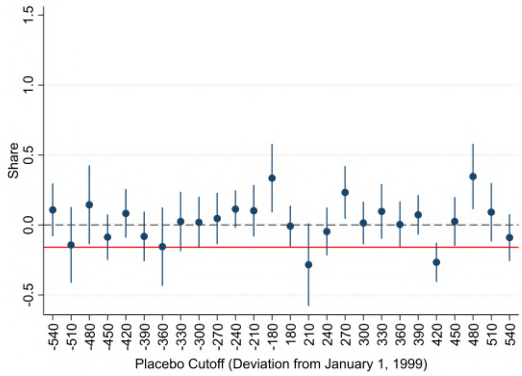
(b) Bandwidth, Charged with a Crime



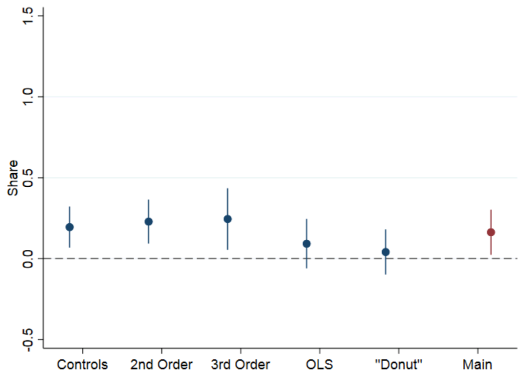
(c) Placebo Cutoffs, Took Any Exam in Lower Secondary School



(d) Placebo Cutoffs, Charged with a Crime



(e) Took Any Exam in Lower Secondary School



(f) Charged with a Crime

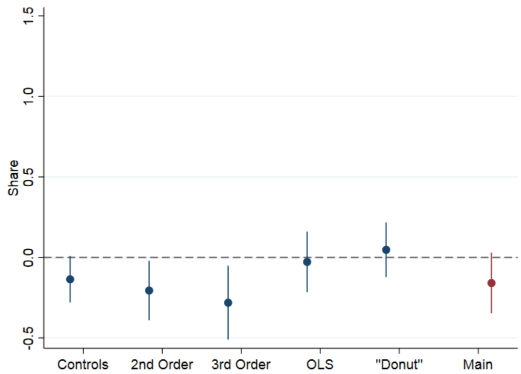


Figure A.VII: Robustness Checks, Boys

Notes: The red dots and bars show the estimate and 95-percent confidence interval based on robust standard errors from column 1 of Table V using the optimal bandwidth from Calonico et al. (2019). The sample is boys born before 2003 using the first parental admission date as assignment to treatment. Panels a and b show the sensitivity to the choice of bandwidth. Panels c and d examine estimates at made-up cutoff points. The red solid line shows the estimates at the true cutoff. Panels e and f compare the main specification (red) to estimates including parental control variables, using 2nd or 3rd order polynomials or OLS, and excluding refugees admitted in the four weeks immediately around the cutoff ("donut"). Control variables are age, age squared, unmarried, female, number of children between 0-2 years old and 3-17 years old, Iraq, Afghanistan, speaks Arabic, Danish 1, 2 or 3 (unknown level is the reference), quota refugee, family-reunification or other refugee (convention refugee is the reference).

B Appendix: Cost-Benefit Analysis of the Language Training Expansion

We calculate the net present value (NPV) of the expansion of language training as:

$$NPV = B - DWL - C = \sum_{i=1}^{18} \frac{b_i - dwl_i - c_i}{(1+r)^i}$$

Where B are the benefits, C are the costs, and r is the societal discount rate. DWL is the deadweight loss (or gain) from changes in the distortion of taxation as a result of the impact on the fiscal budget. We also calculate the benefit-cost ratio, and the time needed before the NPV is positive.

We include gross earnings as the only benefit. They are estimated directly in section **??**. It implies that any positive impacts on the offspring are neglected. The costs of the additional hours provided in the language course is measured by the operating costs described in section IV and taken directly from the law we analyze (Act on Integration, section 45(5), 45(6) and 45(7) and section 59(3) and 59(4)). In addition to these costs and benefits, we include the deadweight loss to society resulting from the impact on the fiscal budget. In the main analysis we assume a tax distortion rate of 50 percent (see e.g. Heckman et al., 2010, for a similar approach).

We calculate the net present value of the change in the fiscal budget as the sum of the changes in tax revenue, savings on welfare benefits and the added costs of the program. Here we assume an average income tax rate of 32 percent for workers and of 26 percent for the unemployed.[50] The savings on welfare benefits are not counted as a benefit in the cost-benefit analysis for society, as it is a transfer (redistribution). All Danish prices are converted to Danish price levels in 2000 using the consumer price index and to USD using an exchange rate of 6.6 DKK/USD.[51]

Hence, in the base scenario (in the first row of Table VI in the main text) we make the following assumptions:

1. The reform extends the previous language learning provision by 30 percent (430 hours) and increases the maximum duration of the program from 18 months to 3 years (see section IV in the main text). We assume the additional language instruction is equal to 6 months (0.3×18 months $=$ 6 months), and assume that the operating costs per month equal the activity-based payment of DKK 3,300 to the municipality. We split the costs equally between the second and third year of

[50]This is based on mean average tax payments for the employed and unemployed in a report by Rambøll and the Danish Center for Social Science Research (2013), the best estimates of the relevant average tax rates we are aware of.

[51]These conversions are applied to all nominal variables throughout the paper.

the program.

2. The discount rate is 3 percent.

3. The tax distortion rate is 50 percent.

We conduct simple sensitivity analyses for each of the items above and add an additional cost through the impact on education (in row two to four of Table VI in the main text):

1. Use the earliest operating costs with direct information on prices per module available (`www.integrationsbarometer.dk`). This is from 2008 and the price is deflated to 2000 prices. The average pay per module in 2008 was DKK 33,735, 20,064 and 17,394 for a participant in Danish course 1, 2 and 3, respectively. We use a weighted average with the share of participants at each course level in 1999 (0.32, 0.35, 0.33; Table I). We assume one module is one sixth of the total hours ($1,830/6 = 340$ hours), so the extra hours amount to 1.4 modules.[52]

2. Discount rate is 7 percent.

3. The tax distortion rate is 0 percent (this is the worst case since it enters as savings in the net costs).

4. Include the costs of additional education for the population below 25 (Appendix Figure A.V).[53] They constitute 13.7 percent of the sample. The price is approximated by the public expenditure per pupil for basic, vocational and academic education. We use the simple average across types of education, and hence not weighted by uptake of immigrants. It is assumed that basic schooling completion takes two years, vocational education takes four years and academic education takes five years. The earliest year for which we have information on prices is 2014 for basic and vocational education and from 2007 for academic education. They are deflated to 2000 prices using the consumer price index. The added cost of education is not included in the base scenario because educational attainment takes time, and we are therefore not seeing all the benefits that comes from it.

We neglect the cost of crime because it is assessed to be very small. This is based on the following: The welfare reduction raised criminal convictions in the first year by 10 percent, arising mainly from an increase in shoplifting. The sentence for shoplifting is a small fine in Denmark, so the costs consist of

[52] The total hours is 1,830 after the reform and there are six modules after the reform. This structure did not exist before. This approach gives as a higher estimate of the operating costs of the program compared to our baseline.

[53] For basic schooling and vocational training: `https://www.uvm.dk/institutioner-og-drift/oekonomi-og-drift/regulerede-institutioner/takstkatalog-og-finanslov/takstkatalog` and for academic education: `https://ufm.dk/uddannelse/videregaende-uddannelse/universiteter/okonomi/uddannelsestilskud` (accessed December 12, 2019).

resources spent by the police, the court, the defence and the prosecution. Using previous estimates of these costs, an estimate of the cost of such a crime is $200. Results of the societal cost-benefit analysis are shown in Table VI and discussed in section VIII.

The impact on the fiscal budget is shown in Table B.I below. Income tax and savings on social welfare payments are positive for the public coffers and the operating costs are the costs for the public coffers.[54]

Although we do not provide estimates of the uncertainty of the numbers in Table B.I, they are all sufficiently large to support the conclusion that the additional investment in language training was beneficial to the society and the government. The net monetary benefit for the refugee is simply the sum of the impact on the difference in post-tax earnings and benefits because the program is free of charge for the participants. This difference is also large and positive, effectively ruling out that the net benefit for the recipient, which could include mental costs of attending the course, is negative.

Table B.I: Fiscal Cost-Benefit Analysis

Government	Income Tax and Welfare Savings (1)	Operating Costs (2)	NPV (3)	B/C (4)	Years Before $NPV > 0$ (5)
Base scenario	18.0	-2.6	15.4	7.0	6
Alternative price	18.0	-3.6	14.4	5.0	6
Discount rate 7%	12.4	-2.4	10.0	5.3	6
With education	18.0	-12.0	6.3	1.5	6

Notes: The rows are the different scenarios described in the text. Columns 1 to 4 are measured in 1,000 USD and column 5 is denoted in years. NPV is the net present value of the investment, B is benefits, and C is costs.

[54] Note a change in the deadweight loss from distortionary taxation is only a cost (or benefit in our case) from the perspective of society and is therefore not part of the fiscal cost-benefit analysis.

Chapter 3

Does Granting Refugee Status to Family-Reunified Women Improve Their Integration?

Does Granting Refugee Status to Family-Reunified Women Improve Their Integration?*

Linea Hasager†

August 29, 2020

Abstract

In many refugee-receiving countries men are the principal asylum applicant, while women are admitted through family-reunification procedures. I document that granting asylum to family-reunified women has significant impacts on economic integration and decreases their risk of being victims of intimate partner violence. Using an event study approach, I find that annual employment and earnings increase by 1.6 percentage points and USD 500, respectively, immediately after asylum recognition relative to family-reunification status. These are large effects compared to the low baseline of virtually zero employment and earnings in the preceding years. At the same time the divorce rate increases by 3.8 percentage points and domestic violence decreases by 0.9 percentage points. The decrease in violence is observed regardless of whether the woman remains married or not, which suggests that the new, more favorable, residence permit improves her bargaining power within the marriage. This is consistent with the predictions from a Nash bargaining model where the risk of being returned to the home country affects the woman's threat point, and thus the allocation of resources within the marriage.

JEL Classification: J12, J15, J61, K37.
Keywords: Refugees, Asylum Recognition, Female Integration, Intimate Partner Violence

*I thank Mette Foged, Giovanni Peri, Jakob Roland Munch, Mia Jørgensen, Mette Rasmussen, Oddbjørn Raaum and Torben Tranæs for helpful comments and discussions. The project benefited from comments from participants at Giovanni Peri's *Ph.D. Advisee Lab*. I gratefully acknowledge support from the Economic Assimilation Research Network (EARN), at the University of Copenhagen, financed by the Innovation Fund Denmark (grant #6149-00024B). All errors are my own.

†Department of Economics, University of Copenhagen, Øster Farimagsgade 5 Building 26, DK 1353 Copenhagen, tlh@econ.ku.dk.

I Introduction

Refugee migration flows exhibit a strong pattern across gender worldwide: Men are more likely to be recognized as refugees, while women are mostly admitted through the family-reunification system.[1] This may put displaced women at a significant disadvantage in their host countries, because their residence permit is contingent on their husband. In case of divorce or death of her partner, the woman will lose her residence permit, which implies that individuals escaping similar conditions in their origin countries face different risks of being returned to their home country simply due to differences in admission status.

At the same time, there exist significant gender differences in labor market integration of immigrants, and in particular for refugees. (Brell, Dustmann, and Preston, 2020) document that immigrants lack behind natives in terms of earnings and employment, and these gaps are most pronounced for refugee women. Moreover, existing evidence from the Nordic countries shows that traditional measures such as welfare benefit reductions and active labor market programs are less effective in pushing female refugees into work (Arendt and Schultz-Nielsen, 2019), while there are significant returns to policies that increase investments in human capital such as language skills and education for immigrant women, see (Arendt et al., 2020). These findings suggest that in order to improve female integration, policies that increase the relative wage of women and their bargaining power within the household may be of particular importance for policy-makers seeking to close the substantial earnings and employment gaps.

Besides the disadvantageous labor market outcomes, family-reunified women also face a higher risk of being subject to intimate partner violence (IPV). Roughly, 2 percent of women, who are family-reunified to a refugee spouse, are victimized by their partner each year in Denmark – and these are only the cases reported to the police. In comparison, 1 percent of other refugee women and 0.2 percent of native women experience intimate partner violence annually. The increased risk of intimate partner violence for family-reunified women may partly reflect the high dependence on husbands, since these women are only allowed to stay in the host country as long as they remain married.

In this paper I study the consequences of recognizing family-reunified women as refugees, which improves their bargaining power within the household and reduces the probability of being returned to the origin country. To estimate the impact of asylum recognition for these women, I adopt an event study approach around the time of asylum recognition. This quasi-experimental approach exploits the

[1](Eurostat, 2020) shows that 62 pct. of first-time asylum seekers in 2019 in the European Union were men. (Statista, 2020) documents that men are more likely to be the principal refugee applicant in the U.S., while women are more likely to be admitted as dependents.

unpredictability of being recognized as a refugee, conditional on application for asylum, by comparing outcomes just around the time of the event. The main assumption behind the empirical strategy is, that in absence of the event occurring, outcomes would have evolved smoothly over time. The sharp changes in outcomes around the event are attributed to asylum recognition, and thus measure the impact of refugee status relative to family-reunification status. Because there is variation in time elapsed between immigration and asylum recognition, it is possible to take account of assimilation profiles and business cycles which may be influencing outcomes.

The results show that divorce rates increase immediately after asylum recognition and stabilize at around 9 percent already a few years later. Simultaneously, the risk of experiencing domestic violence drops significantly, while employment rates, earnings and enrollment in education increase following asylum recognition. These findings are similar in magnitude regardless of whether the women remain married or divorce their partners, but the impact on divorce rates is higher for women from countries with higher female labor force participation. The findings are consistent with a household bargaining model that incorporates violence, bargaining over traditional gender norms and risk of return to the origin country.

My findings shed light on the large gender disparities in refugee admission practices, and document that the type of residence permit has important implications for several key measures of female integration. There exists ample evidence showing that the type of residence permit for immigrants matters for their economic performance and well-being. However, to the best of my knowledge, there exists no evidence on the consequences of the stark gender imbalance in the refugee admission and family-reunification system, which systematically creates unfavorable conditions for displaced women. The contribution of this paper is to compare the differences in outcomes resulting from asylum recognition as opposed to family-reunification status for women.

Previous research has been devoted to estimating the impacts of granting legal status to undocumented immigrants, primarily focused on the United States. These illegal immigrants face a risk of deportation and are not allowed to work in their host country, which may directly impair labor market outcomes. Evidence from Hispanic immigrants in the U.S. shows that legalization increased wages for both men and women, while the results on employment are mixed (Amuedo-Dorantes, Bansak, and Raphael, 2007; Orrenius and Zavodny, 2015).[2] Two recent studies estimate the impact of temporary legalization on mental health and teenage pregnancy in the U.S.[3] They show that legalization of immi-

[2](Amuedo-Dorantes, Bansak, and Raphael, 2007) find that employment rates decreased, while (Orrenius and Zavodny, 2015) find that employment decreased for men, but increased for less-educated women.

[3]Deferred Action for Childhood Arrivals (DACA) provided work authorization and deferral from deportation for two years

grant mothers improved mental health for their children (Hainmueller et al., 2017), while the temporary legalization of Hispanic youths decreased teenage pregnancy rates.[4]

Existing literature has also considered the impacts of obtaining citizenship. (Hainmueller, Hangartner, and Pietrantuono, 2015) document that naturalization fosters political integration of immigrants in Switzerland, and (Hainmueller, Hangartner, and Pietrantuono, 2017) show that naturalization increases social integration, such as membership of local clubs, reading local newspapers etc. However, (Dahl et al., 2020) find that automatic birthright citizenship for children born to immigrants in Germany decreased well-being (life satisfaction and self-esteem) for girls, while improving it for boys.[5]

Most closely related to my work are two working papers considering the distinction between permanent and temporary status for refugees along with a literature on the relationship between intimate partner violence and economic dependence on husbands. (Kilström, Larsen, and Olme, 2018) find that a decrease in the probability of receiving permanent status coupled with an incentive to achieve more labor market attachment and education resulted in increased education enrollment rates and lower earnings for refugee women in Denmark.[6] A similar study in Sweden documents that temporary status coupled with work incentives had no impact on employment, earnings or education, but there were indications of increased participation in language training within the first six months after admission, see (Blomqvist, Thoursie, and Tyrefors, 2018). (Hidrobo, Peterman, and Heise, 2016) show that transfers to urban poor women in Ecuador decreased domestic violence, while (Aizer, 2010) documents that increases in relative wages for women in California have also reduced intimate partner violence by improving female bargaining power.

Thus far, this leaves the question of the consequences of asylum versus family-reunification status unanswered. The contribution of this paper is to document that these different residence permits have important implications for the persons admitted through the two systems. Second, my results provide one explanation for why refugee women lack significantly behind their male counterparts in terms of economic integration. Moreover, I show that the current practice not only hinders female integration, but also puts women at higher risk of intimate partner violence.

The remainder of this paper is structured as follows. In Section II, I present a simple Nash bargaining

for undocumented youths in the U.S. The eligibility requirements incentivized enrollment in and graduation from school.

[4](Hainmueller et al., 2017) show that children whose mothers were eligible for DACA had significantly lower rates of adjustment and anxiety disorders, than children whose mothers were at risk of deportation. (Kuka, Shenhav, and Shih, 2019) show that DACA also reduced teenage pregnancy rates for Hispanic youths, which may be due to an increased education incentive.

[5]The deteriorating well-being of girls is explained by higher parental pressure of conforming with traditional origin country norms when they are offered German citizenship.

[6]Male refugees' outcomes were not affected by the reform, except for a decrease in property crimes.

household framework, which predicts an increase in divorce rates along with decreases in domestic violence and traditional gender norms following a reduction of the woman's risk of return to the origin country. Section III describes the refugee admission and family-reunification system in Denmark, which by and large is governed by the internationally recognized 1951 UN Refugee Convention. Section IV describes the detailed individual-level data on residence permits, asylum applications and admissions along with definitions of the main outcomes in the empirical analysis. In Section V, I present the event study methodology applied in estimating the impacts of asylum recognition for family-reunified women. Section VI shows tests of the validity of the empirical strategy along with the main results. Finally, Section VII concludes the paper.

II Theoretical Considerations

In order to fix ideas, I set up a simple theoretical model which describes the impacts of a reduction in risk of the woman being returned to her home country. The model captures intra-household allocation of consumption, the prevalence of intimate partner violence and the level of traditional gender norms. Following (Aizer, 2010) I set up a Nash bargaining model which incorporates domestic violence. I supplement this model with a risk of being returned to the origin country, which poses a significant disutility to the woman. As another deviation from (Aizer, 2010), I include traditional gender norms in the couple's utility functions. The man's utility is assumed to be increasing in traditional gender norms, while the woman's utility is decreasing in traditional gender norms. Traditional gender norms could include the extent to which the woman participates in the labor market, enrollment in education and fertility choices. The model shows that a decrease in the woman's risk of being returned to her home country leads to a decline in violence and fewer traditional gender norms within the marriage.[7]

A. Nash Bargaining Model with Risk of Being Returned to the Origin Country

Let $U_w(C_w, S, 1-G)$ describe the woman's utility which is increasing in her own consumption (C_w), increasing in safety (S) and decreasing in traditional gender norms (G). The man's utility is denoted by $U_m(C_m, V, G)$. His utility is increasing in his own consumption (C_m), increasing in traditional gender norms (G) and increasing in male autonomy which is achieved through violence (V).[8] There is an upper

[7]A reduction in risk of returning to the origin country also increases the expected value of finding a job for both the job seeker and the employer. I abstract from this in the Nash bargaining model, but (Kilström, Larsen, and Olme, 2018) present a search and match model describing these features.

[8]I assume that the violent behavior increases his self-esteem, power, behavior modification or any other psychological factors which can be present in a violent relationship, see (Aizer, 2010) and (Farmer and Tiefenthaler, 1997).

bound to violence (death) $\bar{V}$, such that $S = \bar{V} - V$. Moreover, the household has total income I, and the woman's share of income, if she was not married, is denoted by α.

If the couple cannot reach an agreement on the intra-household allocation of resources, they both receive the utility of being single, which is denoted by the pair of utilities $(d_m, d_w) = (U_m((1-\alpha)I, 0, 0), U_w(\alpha I, \bar{V}, 1) - \delta R)$. In this case there is no intimate partner violence or imposition of traditional gender norms. The parameter δ captures the woman's risk of being returned to her home country, while R is the disutility she gets in that case. This disutility arises because she has deliberately fled her home country, so she strictly prefers to reside in the host country. This could be due to risk of persecution, war or other undesirable conditions in her origin country.

(Aizer, 2010) shows that this problem constitutes a Nash bargaining problem and that a Nash bargaining solution provides a unique solution to the problem, which is solved by using the asymmetric bargaining solution from (Kalai, 1983). Formally, (U_m^*, U_w^*) maximizes the following expression:

$$\begin{aligned} &(U_m - d_m)^{\tau}(U_w - d_w)^{1-\tau} \\ s.t. \qquad &(U_m, U_w) \geq (d_m, d_w) \\ and \qquad &(U_m, U_w)\ feasible. \end{aligned} \tag{1}$$

Here τ denotes the bargaining power of the man. This framework allows me to analyze the impacts on the level of violence and traditional gender norms following a reduction in the woman's risk of returning to her origin country when she is granted her own asylum status. In Appendix A, I derive the first-order conditions for the optimization problem.

The model predicts a decrease in violence, when the woman's risk of being returned to her origin country is decreased. Because her utility of divorce increases if her risk of being returned to her home country is lower, she demands a more favorable allocation of household resources, which includes less violence. Similarly, the level of traditional gender norms imposed within the marriage will also decrease as a result of a decreased risk of return to the origin country, because her threat point within the marriage has improved. It is possible that the reduction in risk of being returned to the woman's origin country is so large that there is no longer any utility pair that will satisfy the conditions for the spouses remaining married. This implies that both spouses are better off by divorcing each other in that case.

These predictions from the bargaining model outline the hypotheses behind the empirical tests carried out in Section VI, where asylum recognition for women is expected to increase divorce rates, labor market participation and education, while decreasing the prevalence of domestic violence.

III Refugee Admissions in Denmark, 1997-2017

During the period 1997 to 2017 Denmark admitted approximately 60,000 persons aged 18 to 64 as refugees or family-reunified to a refugee spouse. The majority of the refugees, 47 pct., are granted asylum in accordance with the internationally recognized 1951 UN Refugee Convention (UNHCR (2010)). The Danish state also grants asylum to individuals who are not covered by the UN convention, but are at risk of inhumane treatment in their home countries, which amounts to approximately 30 pct. of the admitted refugees ("de facto-" or "B-status"). Another 10 pct. are resettled through the international UN quota system, and the remaining persons are admitted on temporary protection grounds and other humanitarian concerns.

Refugees are secured the right to family-reunification which implies that 22 pct. of the 60,000 adults admitted in 1997 to 2017 arrived through the family-reunification system.[9] A remarkable imbalance emerges when studying the admissions by gender; 80 pct. of the adults admitted through the family-reunification system are women. On the other hand, out of all men who were admitted, 95 pct. were granted refugee status.[10] This reveals a strong pattern in refugee migration, which resembles that of other countries in Europe and the United States (see (Eurostat, 2020) and (Statista, 2020)). The observed pattern may reflect individual preferences and abilities to migrate along with family-based decision-making and gender differences in qualifying as a refugee. Reports from the United Nations show that females are at high risk of being subject to violence during the journey from the origin country to their destination, and it may be difficult bringing children with them.[11] This may explain why more young single men engage in the journey to seek asylum, and why it is common for male husbands to endure the risky journey from refugee-sending countries, and then subsequently apply for family-reunification of wives and children, which secures a safe passage for them to the destination country.

A second contributor to the pattern in admission types may be the grounds for asylum recognition stipulated in the 1951 UN Refugee Convention and the special Danish asylum recognition practice. The UN Convention applies to individuals persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion. For example, individuals who are at risk of being killed by their home country's regime due to political activities, desertion from the army or fleeing military conscription. In addition, the Danish state grants asylum to individuals at risk of torture, death penalty

[9] Adult refugees can have spouses and children reunified to them in the host country. Minors may also have their parents reunified to them. In very special cases, other family members may be reunified to the refugee. In order to have a spouse family-reunified, the couple must prove that they are in a valid marriage based on strict conditions such as previous cohabitation, children etc., see (The Danish Immigration Service, 2020). This hinders the occurrence of pro-forma marriages.

[10] Of all women admitted to Denmark, 60 pct. of them were admitted as refugees.

[11] See (United Nations University, 2015) among others.

or persecution in the home country if the UN Convention does not apply.[12] Political engagement and military conscription is more widespread for men than for women in most refugee-sending countries, which implies that it is harder for women to prove that they are individually persecuted for the reasons described above.[13]

An important feature of the Danish refugee admission system is that the legal residence permit for the family-reunified woman is contingent on her spouse.[14] This means that the family-reunified spouse will lose her permit in case the refugee husband loses his permission to stay, if the couple is divorced or in the event of death of the partner. This poses a significant disadvantage to the group of women admitted through the family-reunification system by increasing uncertainty about residency and shifting intra-household bargaining power in favor of the husband.

One final important aspect is worth noting; while asylum seekers in Denmark were not allowed to work until 2013, both recognized refugees and their family-reunified family members have the right to work, free access to the Danish health care system and are allowed to vote in municipal elections after three consecutive years of residency.

IV Data

The analysis is based on administrative data on the full population in Denmark from 1997 to 2017. I use unique individual identifiers to track individuals over time and across different registers. Furthermore, the data allows me to link individuals to their spouses and children. To study the patterns in refugee migration and changes to legal residence status, I use information from the register on residence permits ("OPHG") which carries detailed information on immigrant status along with dates of application for and approval of residence permits. I link this data to annual outcomes from the Integrated Database for Labor Market Research ("IDA"), the Income Register ("IND"), police registers on victimization, crim-

[12] The Danish government has on some occasions passed special laws granting asylum to ethnic groups who are not individually persecuted. This includes Bosnian refugees in 1995, refugees from Kosovo in 1999, and Syrians in 2015, see (Hvidtfeldt and Schultz-Nielsen, 2017).

[13] See (United Nations, 2020) for the numbers on political representation by gender. Typical refugee-sending countries such as Syria and Iran have mandatory military conscription for men but not for women, see CIA (2020). Somalia has military conscription for both genders, but it has not been applied systematically for women, see UNHCR (1991).

[14] This is similar to other countries such as the United Kingdom and Sweden. According to EU Directive 2003/86/EC, persons who have been family-reunified to a refugee sponsor are not automatically granted an autonomous residence permit in case of divorce or death of the sponsor within the European Union. However, they may apply for an autonomous residence permit depending on national law in each member state. In the U.S. refugees and their derivative refugees (spouses and children) may apply for a permanent residence permit one year after entry. If the principal applicant dies before permanent residence is achieved, the derivative refugee spouse may still apply for adjustment of her status. If the spouses divorce before achieving permanent status, the derivative refugee spouse cannot apply for permanent status based on the former spouse. Instead she must apply for her own asylum, and if she is approved, she can then subsequently become eligible to apply for permanent residence.

inal charges and convictions of crime ("KROF", "KRSI", "KRAF"), enrollment in and completion of education ("UDDA") and the Population Register ("BEF"). This is complemented with macro data from the World Bank and the UN on origin country characteristics such as female labor force participation, fertility and political representation merged to the individual-level data by origin country and year of immigration.[15] The next section describes the sample and variables in more detail.

A. Descriptive Statistics

I study a population of 2,260 adult women (18 years or older) who are initially family-reunified to their refugee spouse, but apply for and obtain their own refugee status subsequently. This amounts to 17 pct. of all women admitted either as refugees or family-reunified to a refugee spouse during the period 1997 to 2016. The women who change to asylum status constitute 37 pct. of all women who were family-reunified to a refugee spouse. The vast majority of these women are granted asylum based on an individual need for protection from persecution in their home countries. Most of them are granted "de facto-status/B-status" (49 pct.) or UN Refugee Convention status (38 pct.), while the remaining 13 pct. are granted asylum due to humanitarian reasons or temporary protection status.

The women are on average 32 years old with two children at immigration, and most have attained a basic level of education upon arrival, see Table I.[16] The majority comes from Iraq (32 pct.), followed by Afghanistan (24 pct.), Syria (16 pct.) and Somalia (10 pct.). There are stark differences between the characteristics of their origin country and Denmark; female labor force participation is generally low (22 pct. compared to 77 pct. in Denmark), the average fertility rate is 3.6 childbirths higher than for Danish women, females hold fewer seats in parliament (14 pct. compared to 37 pct.), many marry at young ages (almost a third is married before age 18), and the average woman spends significantly more time on unpaid domestic chores than men and Danish women (the latter spend 3.8 hours a day on average).[17]

The women whose asylum applications are approved are a selected group compared to women who do not apply for asylum. They also differ compared to the women who are unsuccessful with their applications. Nevertheless, in terms of age, educational attainment and origin country characteristics, they are quite similar to other female refugees and family-reunified women who do not change status (see Table I and Appendix Table B.I). On average the women who change status have more children at immigration, especially compared to women admitted as refugees. This may reflect that the latter are

[15] The macro data are sourced from (The World Bank, 2020) and (United Nations, 2020), respectively.

[16] The information on education acquired abroad is only available for 70 pct. of the sample who were surveyed by a statistical agency.

[17] The numbers for Denmark are from 2018, except for hours spent on domestic work which is measured in 2009.

more likely to be unmarried at immigration.

Most of the women who change residence status are married to a husband of the same nationality as themselves, and the husband generally has attained more education in the home countries than the wife. The husbands of the women who change status also have higher education levels than the husbands of family-reunified women who are not recognized as refugees, see Appendix Table B.I. Their employment rates and earnings at the time of the wife's immigration are below those of husbands to women who are not recognized as refugees. But the levels resemble those of men where both spouses are admitted as refugees.

Figure I illustrates the distribution of immigration, applications and recognition across calendar months. Panel a shows that the distribution of family-reunifications is fairly uniform over the calendar year, with fewest admissions in December (2.3 percentage points below the monthly average), while November is the month with the most admissions (4.6 percentage points above the monthly average). There are slightly more applications for asylum towards the end of the calendar year (Panel b), and these months are also the ones where most of the women are recognized as refugees (Panel c). The latter reflects that the average processing time of an asylum application is just above one year. The bulk of the successful asylum applications are lodged within three months of family-reunificiation, but a few are lodged before and some have applied after several years in Denmark (Panel d). The average wait time for asylum recognition is more than a year, but many cases are processed within 250 days (Panel e), which implies that a good fraction of cases are resolved within two years of family-reunification (Panel f). My empirical analysis is based on variation in the time between family-reunification and asylum recognition depicted in Panel f.

B. Definition of Main Outcomes

The main outcomes of interest are annual gross earnings deflated to 2000-level using the Consumer Price Index (CPI) from Statistics Denmark and converted to USD, and employment measured as a fraction of a full working year.[18] If the worker was a full-time employee during the whole year, this measure takes the value one. If the worker was a part-time employee or not employed in some periods during the year, the fraction is less than one and measures the share of a full-time equivalent. Second, I analyze incidents of intimate partner violence based on cases reported to the police on violence and sexual assault where the woman's husband was charged for the offence.[19] Third, I study investments in human capital by

[18]I use the exchange rate from the Danish Central Bank on March 27, 2019.

[19]This measure likely only captures the most severe cases if less severe cases are not reported to the police.

looking at enrollment in formal education at Danish education institutions.

V Empirical Strategy: Event Study Methodology

The ideal experiment to study the impact of granting women refugee status as opposed to family-reunification would be to randomize residence permits. Such an experiment does not exist because refugee status is granted in accordance with the UN Refugee Convention or similar circumstances. This poses a challenge to identifying the causal impact of one type of residence permit versus another, because the decision to apply for refugee status is likely not orthogonal to individual characteristics. However, given application for asylum, the timing of the event of being recognized as a refugee is unpredictable from the point of view of the asylum seeker. The average waiting time from application to recognition of asylum was more than a year (median = 267 days) during the period studied, see Table I. In fact, the asylum seeker does not know whether or not the application will be successful, so the abrupt event of being granted asylum generates discontinuous changes in labor market performance and other outcomes that would otherwise have evolved smoothly over time. Therefore, the research question posed in this paper is well-suited for applying an event study methodology.[20] The main assumption for identification is, that in absence of the event, the women's outcomes would have evolved smoothly over time. Thus, the sharp changes in outcomes just around the event can be attributed to asylum recognition, and the estimates capture the impact of obtaining refugee status relative to having family-reunification status.

I index all years relative to the event of receiving asylum ($t = 0$) and report annual outcomes four years before and seven years after this event. In the baseline specification I run the following regression:

$$y_{ist} = \sum_{j \neq -1,-4} \beta_j \cdot \mathbb{1}[j = t] + \gamma_s + \alpha_i + \varepsilon_{ist} \tag{2}$$

Here y_{ist} denotes an outcome for individual i in calendar year s at event time t. To capture the dynamic effects, a full set of event time dummies is included (first term on the right hand side). I include calendar year FE γ_s to control nonparametrically for any time trends and business cycles, and α_i denotes individual FE, taking account of the unbalanced panel of individuals. Because there is variation in the number of years elapsed between family-reunification and asylum recognition, the dynamics of the event time dummies are distinct from the assimilation pattern related to the number of years since immigration (family-reunification).

[20]Examples of empirical applications of this approach include (Kleven, Landais, and Søgaard, 2019) and (Dobkin et al., 2018) who study the impacts of childbirth and hospitalizations, respectively.

Following (Borusyak and Jaravel, 2018) I omit two event time dummies at $t = -1$ and $t = -4$, which serves as a test for the existence of a non-linear pre-trend in outcomes. These tests consistently show that there is no pre-trend in the outcomes of interest, see Figure II. Omitting the event time dummy at $t = -1$, implies that the estimated coefficients measure the effect of asylum recognition relative to the year before asylum was granted. I cluster the standard errors by individuals.

VI Results

A. Testing for Pre-Trends

Asylum applications are evaluated based on the individual's need for protection, and from the point of view of the asylum seeker, it is uncertain whether the application will be successful. In addition, it is also unpredictable when asylum will be granted if the application is successful. Figure II depicts the results from Model (2), testing for a pre-trend in the main outcomes of interest. In neither of the five outcomes, the F-test of joint insignificance can reject that the four pre-period dummies are statistically equal to one another. The F-stastistics are small (between 0.48 and 1.71), and their corresponding p-values are between 0.62 and 0.18. This supports the assertion that the timing of asylum recognition is unpredictable for the asylum seekers, and thus they do not alter their behavior in anticipation of asylum recognition in the years leading up to it.

B. Main Results

In the first year immediately after asylum recognition there is a significant 3.8 percentage points increase in divorce rates, which stabilizes at around 9 percent in the following years, see Table II. This may reflect that some of the couples are undergoing a separation period in the first years before they can officially divorce. Simultaneously, the risk of being subject to intimate partner violence drops by 0.9 percentage points, and this risk stabilizes at a significantly lower level throughout the time period. The women also benefit in terms of labor market outcomes. Their employment rate and earnings increase in first year immediately after asylum recognition, and then gradually increase further during the first four years. The levels then subsequently stabilize at roughly 10 percentage points higher annual employment and USD 2,500 higher annual earnings.[21] These are large effects compared to the very low baseline of hardly any employment or earnings during the first years. Average annual employment and earnings in the four

[21]The modest amount is not conditional on employment and thus reflects that a significant part of these women do not find (full-time) employment.

years leading up to asylum recognition are a modest 1.6 percent and USD 600, respectively. Moreover, there are positive effects outside the labor market; enrollment in formal education increases gradually, and the annual enrollment rate is 0.5-0.6 percentage points higher in years 5-6 after asylum recognition. Compared to the baseline of no enrollment in education in the preceding years, this is a substantial increase.

C. Heterogeneity in the Estimated Impacts

It is informative to split the results by marital status and by a measure of gender norms, namely female labor force participation in the origin country. Figure III shows that the reduction in domestic violence prevails both for women who are not divorced (Panel a) and for women who divorce their husbands at some point throughout the period (Panel b).[22] The results in Panel b are more noisy, since only around 10 percent of the sample divorce their spouse. These findings suggests that the observed decrease in intimate partner violence is not entirely driven by increased divorce rates. The decreased probability of return to the woman's home country also affects the level of domestic violence within the marriage by increasing her utility of divorce, and thereby changing her threat point.

The improvements in employment and earnings prospects are also observed regardless of marital status, although the earnings impacts appear to be slightly larger for the women who are divorced.[23] This is possibly explained by the need to become self-supporting when divorcing their husbands.[24]

The impacts on enrollment in education are driven by married women, while the education impacts for divorced women are negative and not statistically different from zero, although subject to large standard errors.[25] This is consistent with the idea, that divorced women are single providers for themselves and their children, implying that they are financially constrained from enrolling in education.

Next, I split the sample by women who come from countries with female labor force participation below and above the sample median. Female labor force participation is measured in the year of family-reunification to Denmark, and it is a measure of the gender norms prevailing in their home countries at the time of emigration. When performing this exercise, my results in Figure IV show that both groups experience a significant increase in divorce rates. However, the impacts are larger for women from

[22]Appendix Figure B.I presents results from formal tests from the following regressions in Panel c: $y_{ist} = \sum_{j\neq -1,-4} \beta_j \cdot \mathbb{1}[j = t] + \sum_{j\neq -1} \kappa_j \mathbb{1}[j = t] \cdot D_i + \eta D_i + \gamma_s + \alpha_i + \varepsilon_{ist}$, where D_i is a dummy for ever divorcing and the remaining variables and parameters are as described in Model (2). The estimated interaction between event time and ever divorcing confirms that the reduction in IPV is similar across the two groups.

[23]The impacts on labor market outcomes are more imprecisely estimated for the group of divorcees, and the two groups are not statistically distinguishable from one another, see Appendix Figure B.I. Panels a, b and d show results from Model (2), where I include D_{ist} as a dummy for being divorced at time t which is interacted with the event time dummies.

[24]A large fraction continue to receive welfare transfers, which can explain the relatively low levels of earnings.

[25]The estimated impacts on the two groups are not statistically different from one another (Appendix Figure B.I).

countries with above median female labor force participation.[26] The effects on domestic violence are similar across the two groups (Figure IV), while the estimated impacts on earnings and employment are significantly larger for women who arrived from countries with higher female labor force participation (Figure V and Appendix Figure B.II). The latter may reflect that women from countries with generally higher female labor force participation arrive with more skills valued in the labor market.

Finally, my results show that women from countries with the lowest labor force participation increase their enrollment in education significantly, but these effects are indistinguishable from the estimated impact on women originating from countries with higher female labor force participation.

D. Robustness Checks and Placebo Tests

I provide evidence that the findings are robust to the inclusion of different fixed effects. Figure VI shows the results from a simple specification, where I control for calendar year fixed effects, but exclude the unit fixed effects from the main specification. This approach controls for any time trends and business cycles without relying solely on within-person variation. Nevertheless, the estimated dynamic effects are remarkably similar to the main results. The estimated impacts on labor income are larger when the individual fixed effects are excluded, because the inclusion of individual effects will demean the variables.

A potential concern is that the cohorts of women immigrating over time are systematically different. One way to address this concern is by controlling for family-reunification cohort fixed effects. This approach compares women who arrived to Denmark in the same calendar year, but have waited a different number of years in Denmark before they were recognized as refugees. Figure VII confirms that there is no pre-trend in any of the outcomes leading up to the event of asylum recognition, while divorce rates, employment, earnings and education enrollment increase significantly following recognition. At the same time incidences of intimate partner violence drop when the women are granted refugee status.

Another important concern is whether the estimated effects can be explained by the typical assimilation pattern observed in many countries.[27] I take account of this by controlling for years since migration in a parametric and nonparametric fashion. In Figure VIII results are shown with calendar year fixed effects along with a second-order polynomial in years since immigration. The estimated coefficients are very similar if I employ a more flexible specification by controlling nonparametrically for years since

[26] Appendix Figure B.II shows regressions tests for the differential impact across the two groups. The countries with above median female labor force participation are Burundi, Eritrea, Kuwait, Myanmar, Rwanda and Serbia (and Somalia for immigration in 1997-1998). The countries with the lowest female labor force participation are Afghanistan, Iran, Iraq, Pakistan and Syria (and Somalia for immigration in the later years of the sample period).

[27] See (Brell, Dustmann, and Preston, 2020) for assimilation profiles of immigrants in several Western countries.

migration (Appendix Figure B.III). The number of years since migration are distinct from the event time dummies, because there is variation in the number of years between family-reunification and asylum recognition. The figure shows a similar pattern for labor market outcomes and the risk of intimate partner violence compared to the main results. This means that the estimated impact of changing status are in excess of the general assimilation pattern observed over time for immigrants.

Next, I show that the estimated pattern for the treated women is not found for the population of women who arrive as refugees nor for the family-reunified women who do not change residence status. I assign these women a random placebo event date of being recognized as a refugee within four years of their actual immigration.[28] Table III presents the estimates for refugee women from regressions identical to the regression equation for the main results in Table II. Table III shows that there are no discontinuous changes in the evolution of outcomes around the placebo event at $t = 0$. Moreover, employment and earnings (Columns (3) and (4)) decrease 2-3 years after the placebo event, which is the opposite pattern compared to the main results for the treated women. Table IV shows the same placebo tests for family-reunified women who are not recognized as refugees. These estimates also demonstrate that there are no discontinuities at $t = 0$, while employment and earnings decrease over time. There is a clear trend in divorce rates, which increase over time for women who do not change status.[29] This is dissimilar to the flat pre-trend and subsequent discontinuous increase to a constantly higher level for the women who change status.

VII Conclusions

I study the consequences of granting female refugees more favorable residence permits. In particular, I estimate the impact of recognizing women, who are initially admitted through family-reunification procedures, as refugees themselves. This improves their bargaining power within the household, because their threat point increases. When they are recognized as refugees themselves, they are able to divorce their husbands without automatically being returned to their origin countries.

By exploiting the unpredictable timing of asylum recognition using an event study approach, I show that divorce rates increase significantly immediately following asylum recognition. The divorce rate stabilizes at around 9 percent in the following years. This effect is particularly large for women from origin countries where female labor force participation is above the sample median.

[28] I use a chi-square distribution with three degrees of freedom for placebo application dates within one year of immigration. Second, I assign placebo recognition dates using a chi-square distribution with three degrees of freedom from the placebo application to three years after this date.

[29] Recall that the divorce rate is normalized at $t = 0$.

In addition, I document that the risk of being subject to intimate partner violence decreases sharply following asylum recognition. This effect is observed both for women who remain married and women who divorce their husbands throughout the period. This suggests that the more favorable residence permit improves the woman's bargaining power within the marriage.

Asylum recognition also has positive consequences for females' employment and earnings trajectories. Employment and earnings increase gradually in the first 4 years after asylum is granted relative to the years prior, where they have family-reunification status. After 4-5 years, employment and earnings then stabilize at significantly higher levels. Finally, I show that human capital investments increase when the women are granted refugee status, as they begin to enroll in the Danish education system.

These patterns are robust to including various sets of fixed effects, and they are consistent with the predictions from a Nash bargaining model incorporating bargaining over violence, gender norms and income within the household.

My work documents a stark gender imbalance inherent in the refugee admission system, which contributes to significant gender inequality among a disadvantaged group of immigrants. Thus, my findings shed light on one of the reasons why female immigrants are lacking behind economically in their host countries. These results have important implications for policy-makers interested in designing immigration policies that can improve economic and social integration of displaced women. A potential way of improving female integration is to automatically process all applications for family-reunification to refugees as asylum applications.

VIII Tables and Graphs

Table I: Summary Statistics for the Population of Refugee Women

	Mean	S.D.
Characteristics at Immigration		
Age	32.66	8.26
Number of Children	2.01	1.97
Days Between Family-Reunification and Application	92.09	379.32
Days Between Application and Asylum	387.92	427.28
Year of Family-Reunification	2003.69	7.05
Year of Asylum Application	2003.91	7.22
Year of Asylum	2004.98	7.37
Education Surveyed		
Basic Education	0.76	0.43
Vocational Education	0.09	0.29
Academic Education	0.15	0.36
Education Not Surveyed	0.30	0.46
Origin Country		
Iraq	0.32	0.47
Afghanistan	0.24	0.42
Syria	0.16	0.36
Somalia	0.10	0.31
Origin Country Characteristics		
Female Labor Force Participation	0.22	0.20
Male Labor Force Participation	0.79	0.03
Fertility Rate	5.32	1.79
Share of Seats in Parliament Held by Women	0.14	0.09
Share of Women Married Before Age 15	0.07	0.03
Share of Women Married Before Age 18	0.29	0.10
Avg. Daily Hours of Domestic Work (Women)	4.88	0.37
Avg. Daily Hours of Domestic Work (Men)	1.35	0.40
Characteristics of Husband at Immigration		
Same Origin Country	0.92	0.26
Employment Rate	0.09	0.24
Any Employment	0.16	0.36
Labor Income (1,000 USD)	3.99	9.79
Husband's Education Surveyed		
Basic Education	0.50	0.50
Vocational Education	0.16	0.37
Academic Education	0.33	0.47
Education Not Surveyed	0.27	0.45
N	2,260	

Notes: Summary statistics for the population of females who were family-reunified to a refugee spouse and subsequently obtained their own refugee status. Age and number of children are measured in the year of family-reunification. Educational attainment acquired abroad shows the distribution across different education levels for those who were surveyed about this. Origin country characteristics are measured at the country level; Labor force participation, fertility rates and the share of seats in parliament held by women are measured in the year of immigration. The share of women married before age 15/18 and average daily hours of domestic work are measured as country averages across all years available in the UN Gender Statistics and Indicators database. The characteristics of the husband are measured in the year of the wife's immigration.

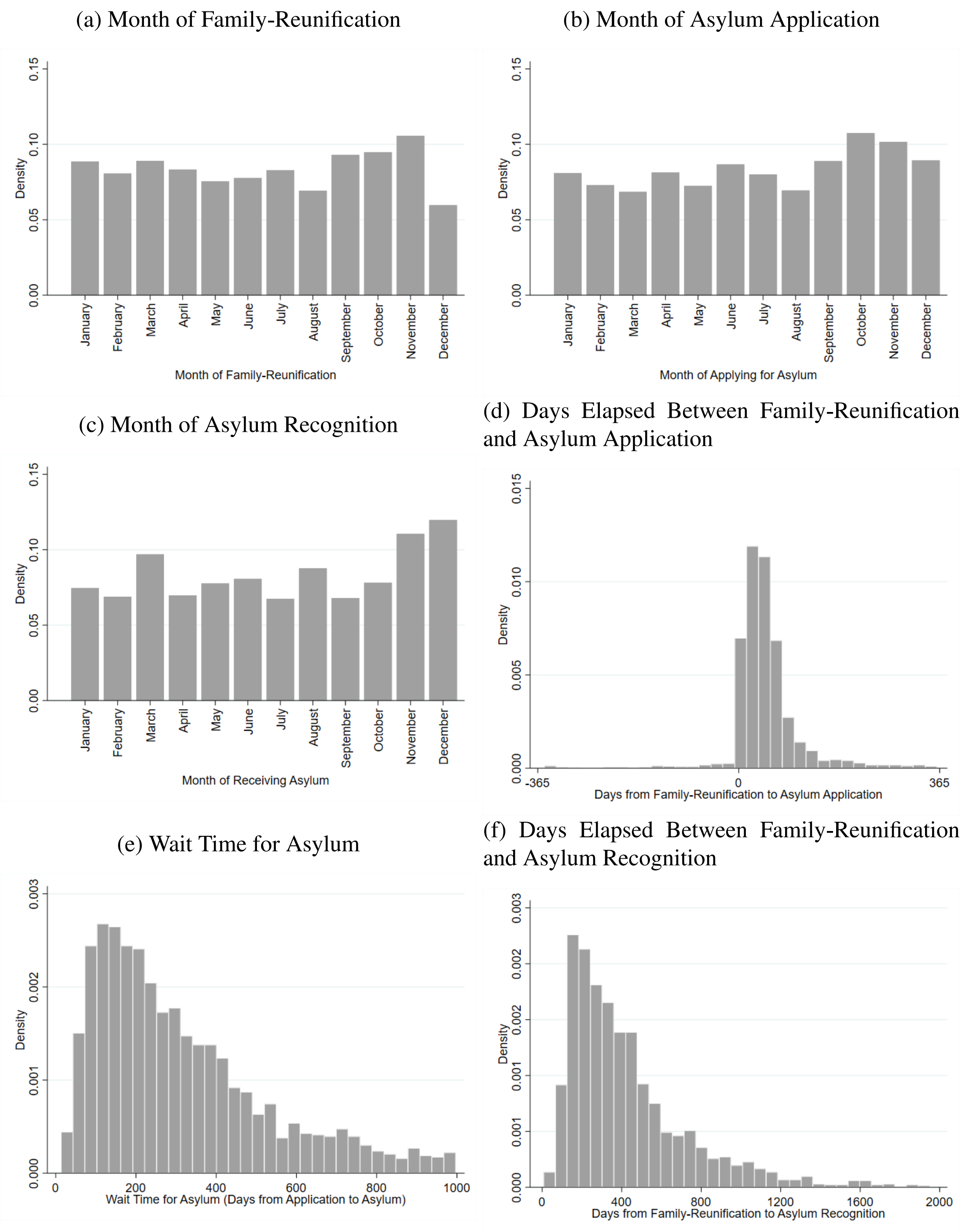
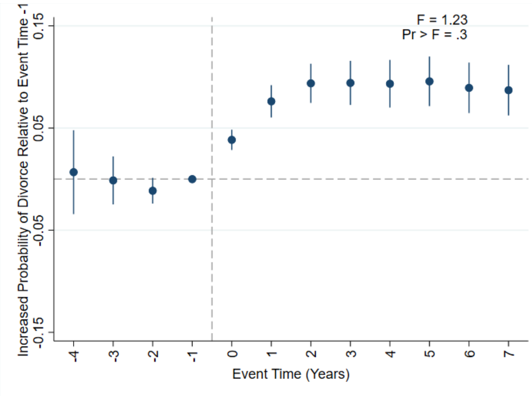


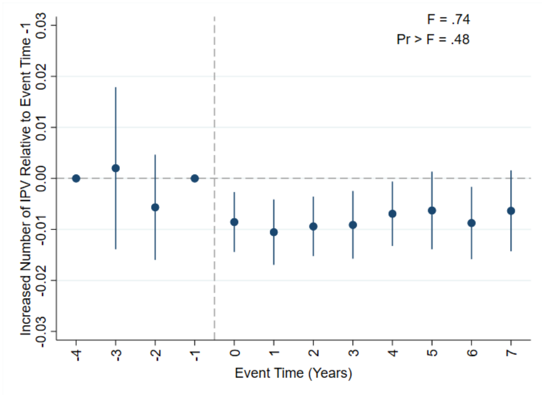
Figure I: Timing of Family-Reunification, Asylum Application and Recognition of Refugee Status

Notes: Panels a-c show the distribution of family-reunification, asylum application and asylum recognition across calendar months, respectively, for the analysis sample. Panels d-f show the distribution of days elapsed between family-reunification and asylum application, days from asylum application to recognition and days from family-reunification to asylum recognition, respectively, for the same sample of women.

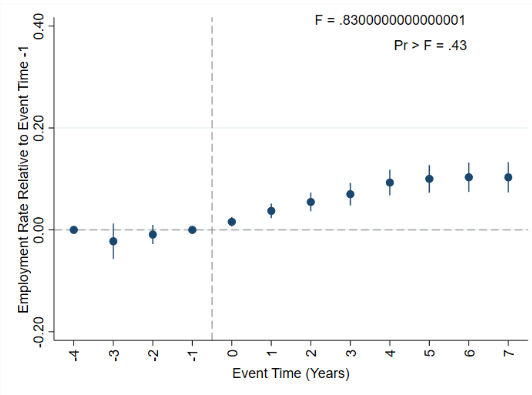
(a) Divorce



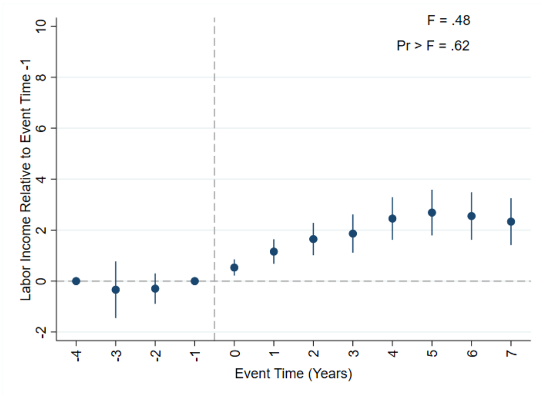
(b) Intimate Partner Violence



(c) Employment Rate



(d) Labor Income



(e) Enrolled in Education

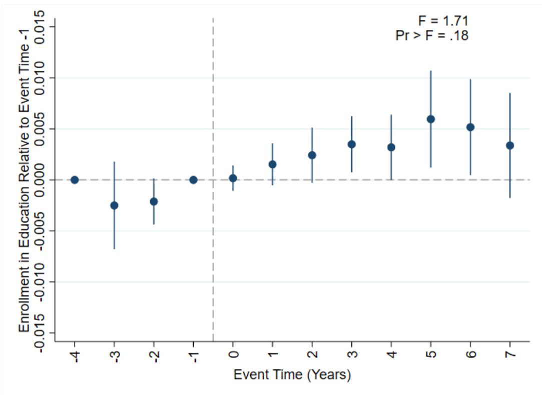


Figure II: Main Results

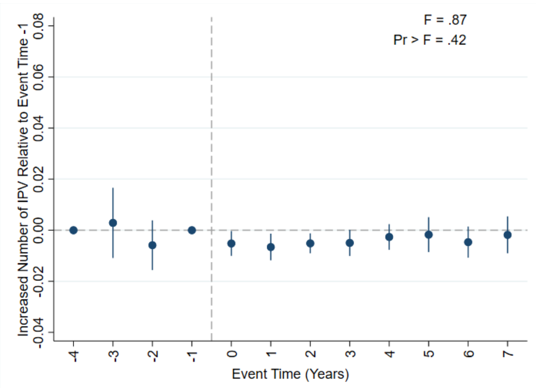
Notes: The blue vertical lines represent 95 pct. confidence intervals based on robust standard errors clustered by individuals. In Panel a calendar year FE are included, but unit fixed effects are not included due to lack of variation in the outcome variable: the majority of women only change marital status once throughout the period if they divorce. The omission of unit fixed effects implies that it is unnecessary to omit the event time dummy at $t = -4$ in Panel a. In Panels b-e calendar year FE and individual FE are included. The F-statistic and the p-value for the joint insignificance of the pre-event dummies are shown in the top right corner of each panel.

Table II: Main Results

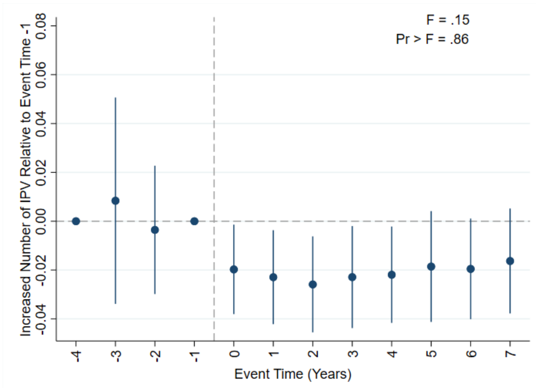
	(1) Divorce	(2) IPV	(3) Employment Rate	(4) Labor Income	(5) Enrolled in Education
-4	0.007	0.000	0.000	0.000	0.000
	(0.021)	(.)	(.)	(.)	(.)
-3	-0.001	0.002	-0.022	-0.338	-0.003
	(0.012)	(0.008)	(0.018)	(0.567)	(0.002)
-2	-0.011	-0.006	-0.009	-0.295	-0.002
	(0.006)	(0.005)	(0.010)	(0.304)	(0.001)
0	0.038***	-0.009***	0.016***	0.532***	0.000
	(0.005)	(0.003)	(0.005)	(0.163)	(0.001)
1	0.076***	-0.011***	0.037***	1.158***	0.002
	(0.008)	(0.003)	(0.007)	(0.246)	(0.001)
2	0.094***	-0.009***	0.055***	1.649***	0.002
	(0.010)	(0.003)	(0.009)	(0.324)	(0.001)
3	0.094***	-0.009***	0.070***	1.864***	0.003**
	(0.011)	(0.003)	(0.011)	(0.383)	(0.001)
4	0.093***	-0.007**	0.093***	2.455***	0.003
	(0.012)	(0.003)	(0.013)	(0.427)	(0.002)
5	0.096***	-0.006	0.100***	2.689***	0.006**
	(0.012)	(0.004)	(0.014)	(0.457)	(0.002)
6	0.089***	-0.009**	0.103***	2.553***	0.005**
	(0.013)	(0.004)	(0.015)	(0.477)	(0.002)
7	0.087***	-0.006	0.103***	2.334***	0.003
	(0.013)	(0.004)	(0.015)	(0.470)	(0.003)
Pre-Asylum Recognition Mean	0.021	0.008	0.016	0.565	0.000
F-Statistic	1.23	0.74	0.83	0.48	1.71
Pr > F	0.30	0.48	0.43	0.62	0.18
N	26,024	26,024	24,443	26,024	26,024

Notes: Robust standard errors clustered by individuals in parentheses. $^{**}p < 0.05, ^{***}p < 0.01$. In all columns the coefficients denote the estimated impact relative to $t = -1$. In columns (2)-(5) the event time dummy at $t = -4$ is also omitted. In column (1) calendar year FE are included, but unit fixed effects are not included due to lack of variation in the outcome variable: the majority of women only change marital status once throughout the period if they divorce. The omission of unit fixed effects implies that it is unnecessary to omit the event time dummy at $t = -4$. In columns (2)-(5) calendar year FE and individual FE are included. The F-statistic and the p-value for the joint insignificance of the pre-event dummies are shown in the bottom of the table. The pre-asylum recognition mean denotes the mean of the outcome in years $t \in [-4, -1]$. IPV measures the annual number of incidents of intimate partner violence. Employment rate is the annual rate of full-time employment. Labor income is measured in USD 1,000 (2000-level).

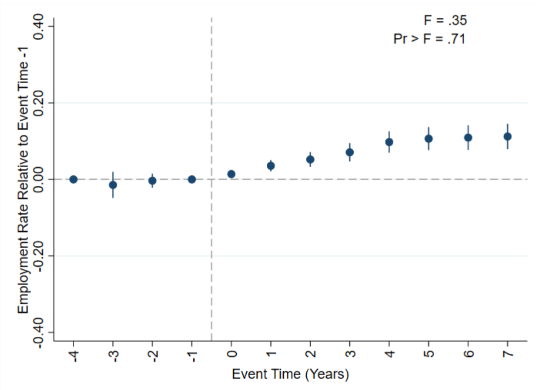
(a) Intimate Partner Violence, Not Divorced



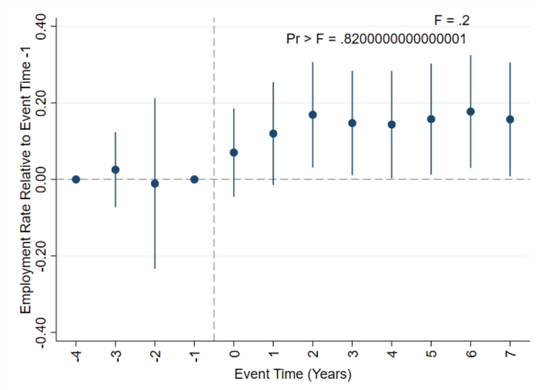
(b) Intimate Partner Violence, Ever Divorced



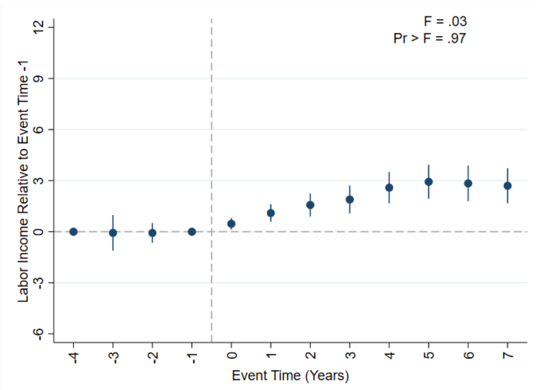
(c) Employment Rate, Not Divorced



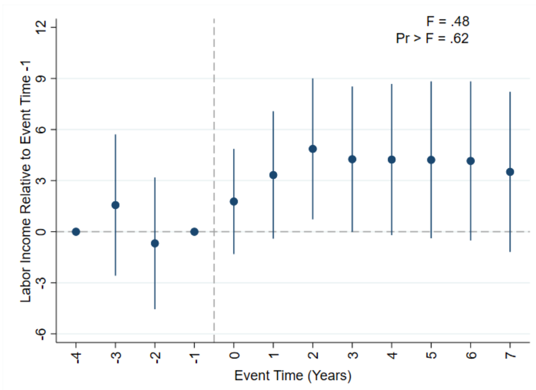
(d) Employment Rate, Divorced



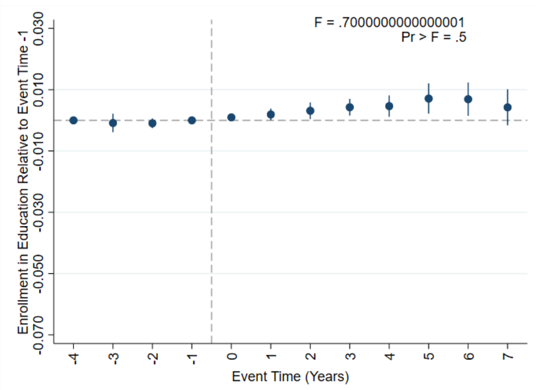
(e) Labor Income, Not Divorced



(f) Labor Income, Divorced



(g) Enrolled in Education, Not Divorced



(h) Enrolled in Education, Divorced

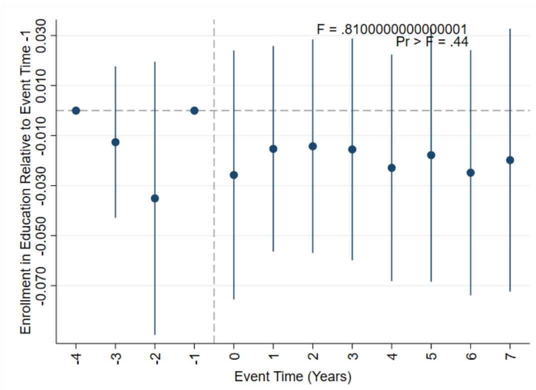
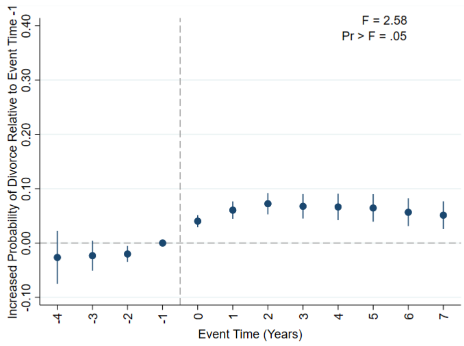


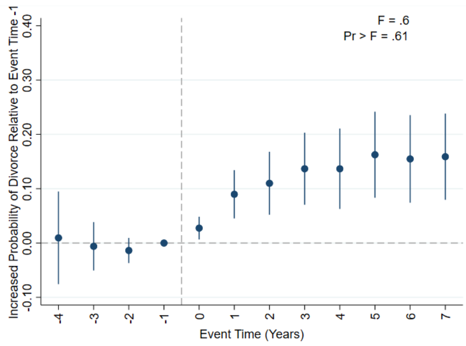
Figure III: Results by Marital Status

Notes: The blue vertical lines represent 95 pct. confidence intervals based on robust standard errors clustered by individuals. In Panels a-h calendar year FE and individual FE are included. In Panel b I study women who ever divorce during the sample period, because she can only be victimized by her husband if the she is married. In the other panels only individuals who are divorced at a given event time contribute to the estimated impact at that event time.

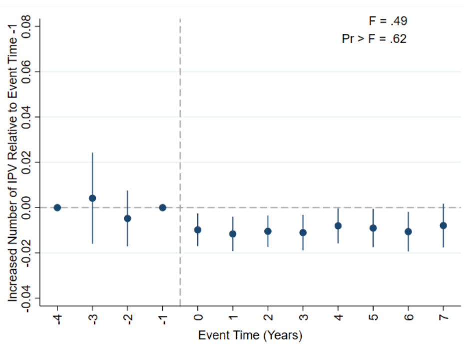
(a) Divorce, Below Median LFP



(b) Divorce, Above Median LFP



(c) Intimate Partner Violence, Below Median LFP



(d) Intimate Partner Violence, Above Median LFP

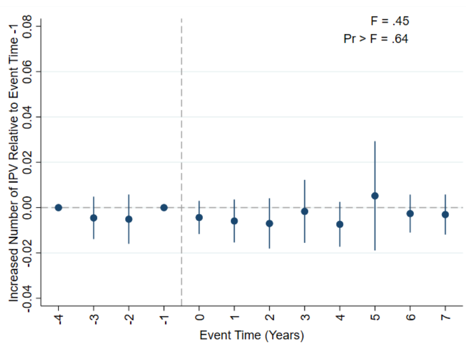
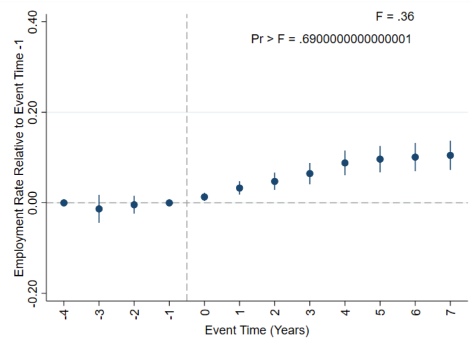


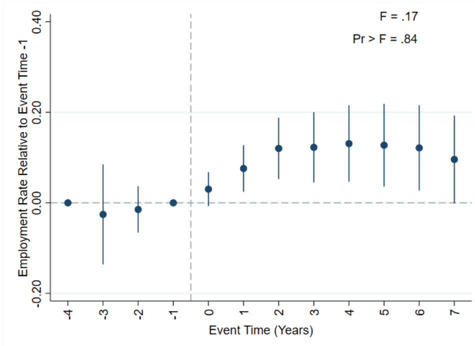
Figure IV: Results by Female Labor Force Participation in Origin Country

Notes: The blue vertical lines represent 95 pct. confidence intervals based on robust standard errors clustered by individuals. In Panels a-b calendar year FE are included, but unit fixed effects are not included due to lack of variation in the outcome variable: the majority of women only change marital status once throughout the period if they divorce. The omission of unit fixed effects implies that it is unnecessary to omit the event time dummy at $t = -4$ in Panel a. In Panels c-d calendar year FE and individual FE are included.

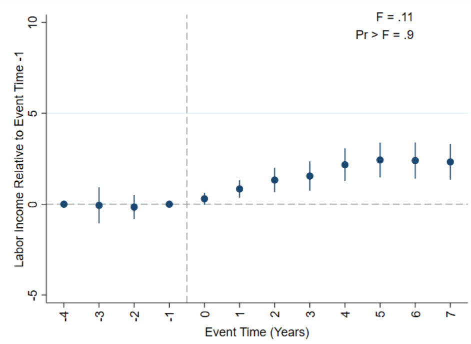
(a) Employment Rate, Below Median LFP



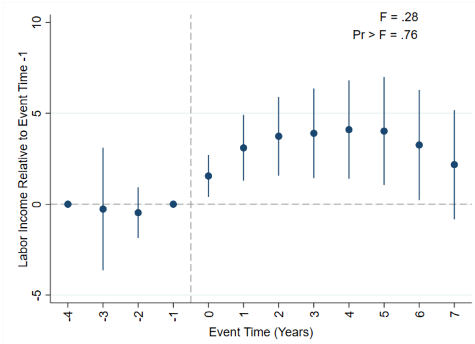
(b) Employment Rate, Above Median LFP



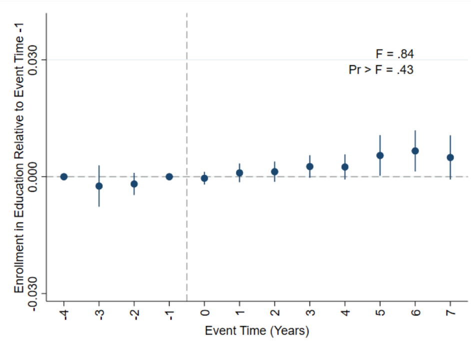
(c) Labor Income, Below Median LFP



(d) Labor Income, Above Median LFP



(e) Enrolled in Education, Below Median LFP



(f) Enrolled in Education, Above Median LFP

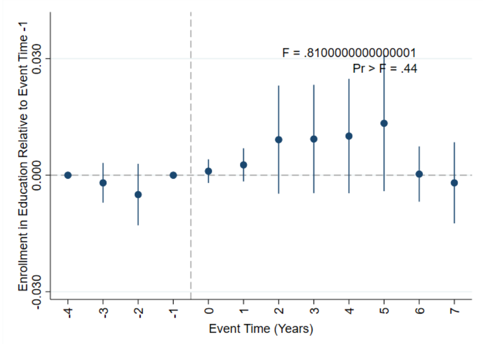


Figure V: Results by Female Labor Force Participation in Origin Country

Notes: The blue vertical lines represent 95 pct. confidence intervals based on robust standard errors clustered by individuals. In Panels a-f calendar year FE and individual FE are included.

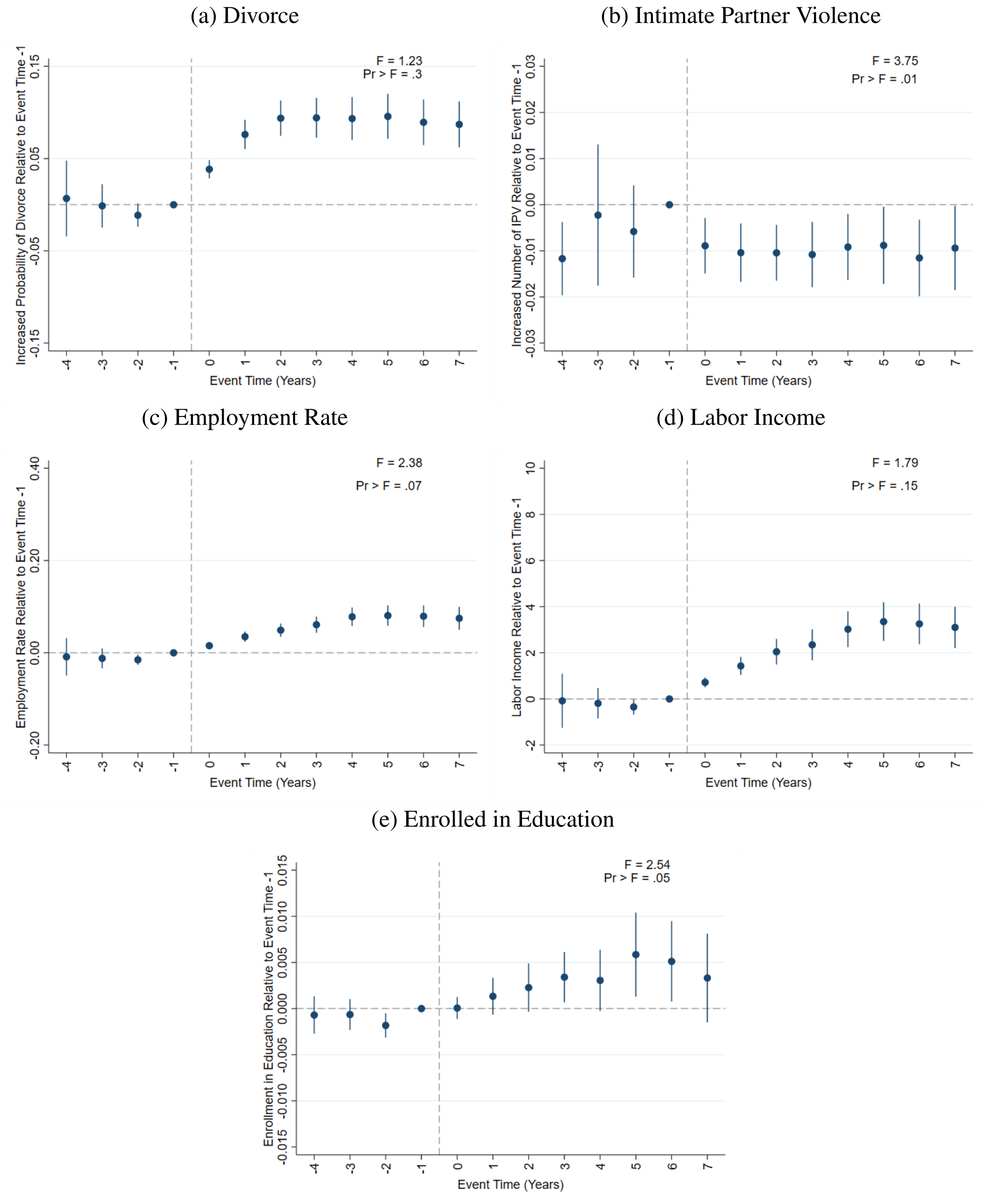


Figure VI: Results With Calendar Year FE

Notes: The blue vertical lines represent 95 pct. confidence intervals based on robust standard errors clustered by individuals. In all panels calendar year FE are included. The F-statistic and the p-value for the joint insignificance of the pre-event dummies are shown in the top right corner of each panel.

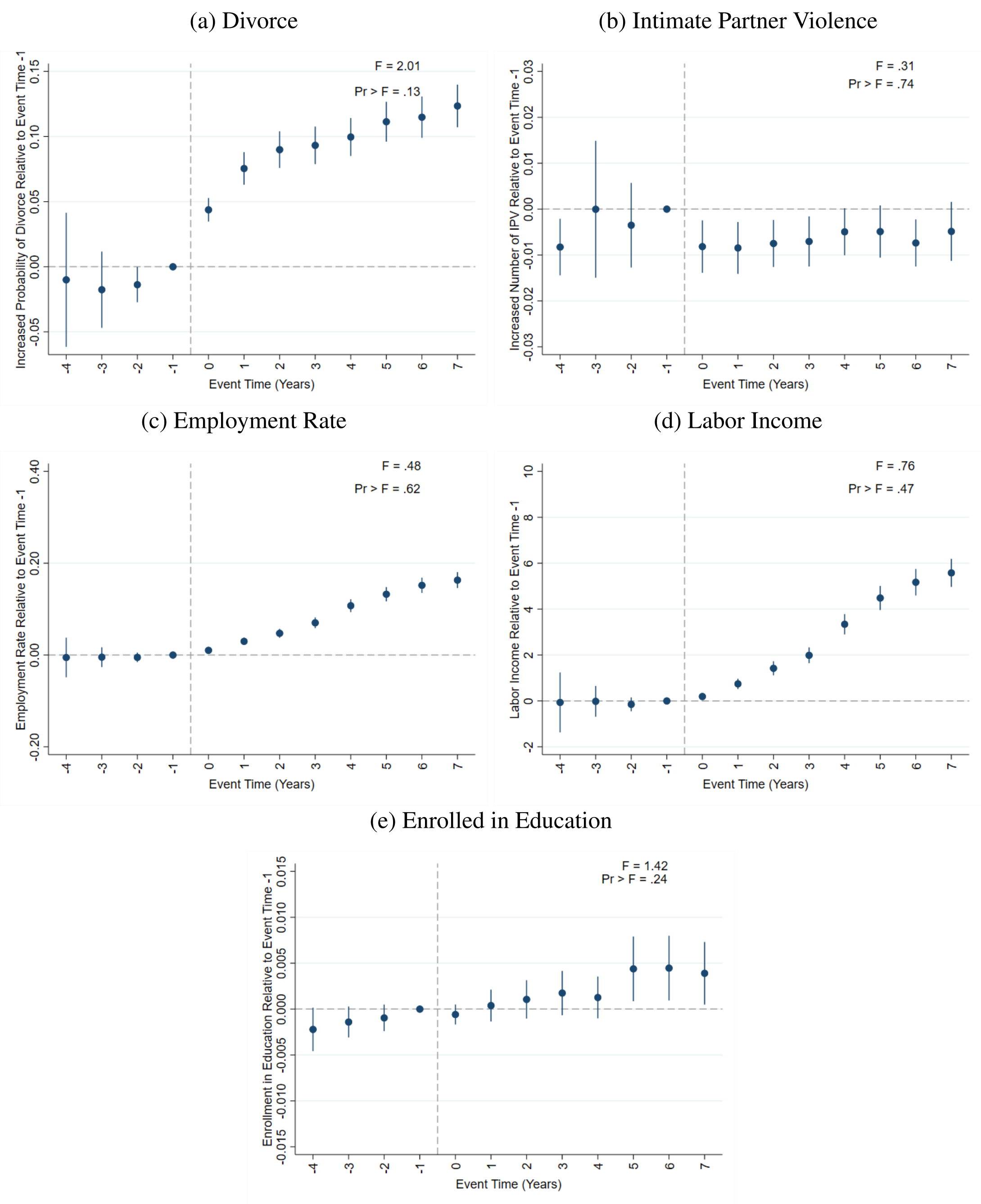


Figure VII: Results With Cohort FE

Notes: The blue vertical lines represent 95 pct. confidence intervals based on robust standard errors clustered by individuals. In all panels family-reunification cohort FE are included (year of first immigration). The F-statistic and the p-value for the joint insignificance of the pre-event dummies are shown in the top right corner of each panel.

(a) Intimate Partner Violence (b) Employment Rate (c) Labor Income (d) Enrolled in Education

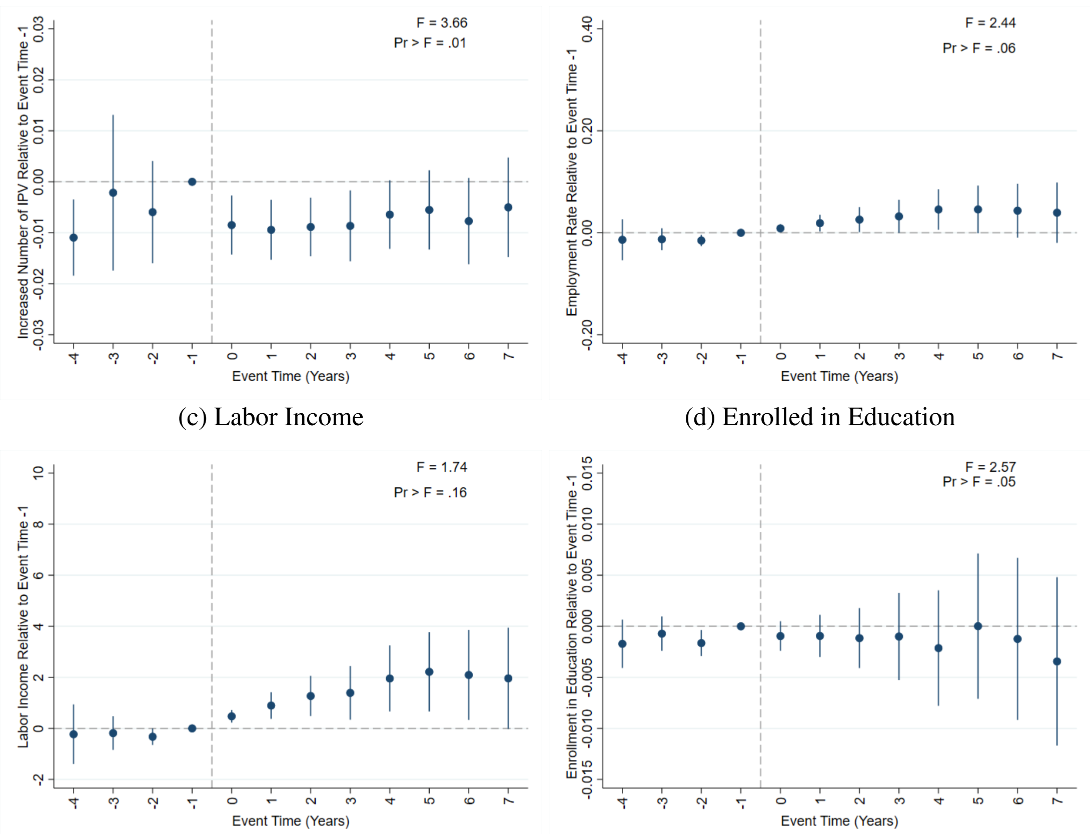


Figure VIII: Results With Calendar Year FE and Years Since Migration

Notes: The blue vertical lines represent 95 pct. confidence intervals based on robust standard errors clustered by individuals. In all panels calendar year FE and a second-order polynomial in years since migration are included. The F-statistic and the p-value for the joint insignificance of the pre-event dummies are shown in the top right corner of each panel.

Table III: Placebo Tests for Women with Refugee Status

	(1) Divorce	(2) IPV	(3) Employment Rate	(4) Labor Income	(5) Enrolled in Education
-4	-0.019***				
	(0.003)				
-3	-0.015***	-0.000	-0.002	-0.051	0.001
	(0.002)	(0.001)	(0.002)	(0.071)	(0.001)
-2	-0.008***	0.001	-0.003	-0.038	0.001
	(0.002)	(0.001)	(0.002)	(0.071)	(0.001)
0	0.010***	-0.000	0.001	0.120	0.001
	(0.002)	(0.001)	(0.003)	(0.104)	(0.001)
1	0.020***	0.000	-0.005	-0.001	0.002
	(0.003)	(0.002)	(0.005)	(0.167)	(0.002)
2	0.029***	-0.001	-0.019***	-0.340	0.001
	(0.003)	(0.002)	(0.007)	(0.220)	(0.002)
3	0.039***	-0.001	-0.031***	-0.734***	0.000
	(0.004)	(0.002)	(0.009)	(0.274)	(0.003)
4	0.046***	-0.000	-0.040***	-1.113***	-0.001
	(0.004)	(0.003)	(0.010)	(0.327)	(0.003)
5	0.051***	-0.001	-0.063***	-1.887***	-0.002
	(0.005)	(0.003)	(0.012)	(0.379)	(0.003)
6	0.060***	-0.000	-0.086***	-2.472***	-0.000
	(0.005)	(0.003)	(0.013)	(0.434)	(0.004)
7	0.068***	-0.001	-0.107***	-3.177***	-0.002
	(0.006)	(0.003)	(0.015)	(0.487)	(0.004)
Calendar Year FE	Yes	Yes	Yes	Yes	Yes
Unit FE	No	Yes	Yes	Yes	Yes
N	88,650	88,562	77,140	88,562	88,562

Notes: Robust standard errors clustered by individuals in parentheses. $^{**}p < 0.05, ^{***}p < 0.01$. The coefficients denote the estimated impact relative to $t = -1$ of an assigned placebo event for refugee women who do not change residence permit. In all columns the event time dummy at $t = -4$ is also omitted, while calendar year FE and individual FE are included. IPV measures the annual number of incidents of intimate partner violence. Employment rate is the annual rate of full-time employment. Labor income is measured in USD 1,000 (2000-level).

Table IV: Placebo Tests for Women with Family-Reunification Status

	(1) Divorce	(2) IPV	(3) Employment Rate	(4) Labor Income	(5) Enrolled in Education
-4	-0.008***				
	(0.003)				
-3	-0.007***	0.000	0.001	-0.096	-0.001
	(0.002)	(0.001)	(0.003)	(0.101)	(0.001)
-2	-0.006***	-0.000	0.003	-0.023	-0.000
	(0.002)	(0.001)	(0.003)	(0.092)	(0.001)
0	0.009***	0.001	0.007	0.028	0.001
	(0.002)	(0.001)	(0.005)	(0.144)	(0.002)
1	0.020***	-0.000	0.006	0.056	0.002
	(0.003)	(0.002)	(0.007)	(0.225)	(0.002)
2	0.029***	0.001	-0.001	0.044	0.003
	(0.004)	(0.002)	(0.009)	(0.293)	(0.003)
3	0.038***	0.001	-0.004	-0.216	0.002
	(0.005)	(0.003)	(0.011)	(0.367)	(0.003)
4	0.051***	0.002	-0.017	-0.535	0.000
	(0.006)	(0.003)	(0.013)	(0.436)	(0.003)
5	0.063***	0.002	-0.028	-1.202**	0.004
	(0.007)	(0.004)	(0.016)	(0.497)	(0.004)
6	0.078***	0.002	-0.040**	-1.616***	-0.003
	(0.008)	(0.004)	(0.017)	(0.565)	(0.004)
7	0.086***	0.003	-0.048**	-2.000***	-0.004
	(0.008)	(0.005)	(0.019)	(0.626)	(0.005)
Calendar Year FE	Yes	Yes	Yes	Yes	Yes
Unit FE	No	Yes	Yes	Yes	Yes
N	47,545	47,492	41,049	47,492	47,492

Notes: Robust standard errors clustered by individuals in parentheses. $^{**}p < 0.05, ^{***}p < 0.01$. The coefficients denote the estimated impact relative to $t = -1$ of an assigned placebo event for family-reunified women who do not change residence permit. In all columns the event time dummy at $t = -4$ is also omitted, while calendar year FE and individual FE are included. IPV measures the annual number of incidents of intimate partner violence. Employment rate is the annual rate of full-time employment. Labor income is measured in USD 1,000 (2000-level).

References

Aizer, Anna. 2010. "The Gender Wage Gap and Domestic Violence." *American Economic Review* 100 (4):1847–59.

Amuedo-Dorantes, Catalina, Cynthia Bansak, and Steven Raphael. 2007. "Gender Differences in the Labor Market: Impact of IRCA." *American Economic Review: Papers and Proceedings* 97 (2):412–416.

Arendt, Jacob Nielsen, Iben Bolvig, Mette Foged, Linea Hasager, and Giovanni Peri. 2020. "Integrating Refugees: Language Training or Work-First Incentives?" *NBER Working Paper No. 26834* .

Arendt, Jacob Nielsen and Marie Louise Schultz-Nielsen. 2019. "Employment Effects of Welfare Policies for Non-Western Immigrant Women." In *Integrating Immigrants into the Nordic Labour Markets*, edited by Lars Calmfors and Nora Sánchez Gassen. The Nordic Council of Ministers, 159–184.

Blomqvist, Niklas, Peter Skogman Thoursie, and Björn Tyrefors. 2018. "Restricting Residence Permits: Short-Run Evidence From A Swedish Reform." *Working Paper* .

Borusyak, Kirill and Xavier Jaravel. 2018. "Revisiting Event Study Designs." *Available at SSRN 2826228* .

Brell, Courtney, Christian Dustmann, and Ian Preston. 2020. "The Labor Market Integration of Refugee Migrants in High-Income Countries." *Journal of Economic Perspectives* 34 (1):94–121.

Central Intelligence Agency. 2020. "Military Service Age and Obligation." URL `https://www.cia.gov/library/publications/the-world-factbook/fields/333.html`.

Dahl, Gordon B, Cristina Felfe, Paul Frijters, and Helmut Rainer. 2020. "Caught between Cultures: Unintended Consequences of Improving Opportunity for Immigrant Girls." *NBER Working Paper No. 26674* .

Dobkin, Carlos, Amy Finkelstein, Raymond Kluender, and Matthew J Notowidigdo. 2018. "The Economic Consequences of Hospital Admissions." *American Economic Review* 108 (2):308–52.

Eurostat. 2020. "Asylum Statistics." URL `https://ec.europa.eu/eurostat/statistics-explained/index.php/Asylum_statistics#Age_and_gender_of_first-time_applicants`.

Farmer, Amy and Jill Tiefenthaler. 1997. "An Economic Analysis of Domestic Violence." *Review of Social Economy* 55 (3):337–358.

Hainmueller, Jens, Dominik Hangartner, and Giuseppe Pietrantuono. 2015. "Naturalization Fosters the Long-Term Political Integration of Immigrants." *Proceedings of the National Academy of Sciences* 112 (41):12651–12656.

———. 2017. "Catalyst or Crown: Does Naturalization Promote the Long-Term Social Integration of Immigrants?" *American Political Science Review* 111 (2):256–276.

Hainmueller, Jens, Duncan Lawrence, Linna Martén, Bernard Black, Lucila Figueroa, Michael Hotard, Tomás R Jiménez, Fernando Mendoza, Maria I Rodriguez, Jonas J Swartz, and David D Laitin. 2017. "Protecting Unauthorized Immigrant Mothers Improves Their Children's Mental Health." *Science* 357 (6355):1041–1044.

Hidrobo, Melissa, Amber Peterman, and Lori Heise. 2016. "The Effect of Cash, Vouchers, and Food Transfers on Intimate Partner Violence: Evidence from a Randomized Experiment in Northern Ecuador." *American Economic Journal: Applied Economics* 8 (3):284–303.

Hvidtfeldt, Camilla and Marie Louise Schultz-Nielsen. 2017. "Flygtninge og asylansøgere i Danmark 1992-2016. Antal, ventetid, bosætning og lovgivning." *The Rockwool Foundation's Research Unit. Working Paper* (50).

Kalai, Ehud. 1983. "Solutions to the Bargaining Problem." *Social Goals and Social Organization: Essays in Memory of Elisha Pazner. Discussion Paper* (556).

Kilström, Mathilda, Birthe Larsen, and Elisabet Olme. 2018. "Should I stay or Must I Go?" *Working Paper 1-2018* .

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2019. "Children and Gender Inequality: Evidence from Denmark." *American Economic Journal: Applied Economics* 11 (4):181–209.

Kuka, Elira, Na'ama Shenhav, and Kevin Shih. 2019. "A Reason to Wait: The Effect of Legal Status on Teen Pregnancy." *AEA Papers and Proceedings* 109:213–17.

Orrenius, Pia M and Madeline Zavodny. 2015. "The Impact of Temporary Protected Status on Immigrants' Labor Market Outcomes." *American Economic Review: Papers and Proceedings* 105 (5):576–80.

Statista. 2020. "Number of Refugees Arriving in the United States in 2018, by Gender." URL https://www.statista.com/statistics/247062/number-of-refugees-arriving-in-the-us-by-gender/.

The Danish Immigration Service. 2020. "Apply for Family Reunification as Spouse/Partner of a Refugee in Denmark." URL https://nyidanmark.dk/en-GB/Applying/Familie/Familiesammenforing/Aegtefaelle%20eller%20fast%20samlever%20til%20flygtning%20i%20Danmark?anchor=canyouapply.

The World Bank. 2020. "Data Bank World Development Indicators." URL https://databank.worldbank.org/source/world-development-indicators.

United Nations. 2020. "Minimum Set of Gender Indicators." URL https://genderstats.un.org/#/downloads.

United Nations High Commissioner for Refugees. 1991. "Somalia: Information on Forced Military Service for Women." URL https://www.refworld.org/docid/3ae6aabfac.html.

United Nations Human Rights Council. 2010. "Convention and Protocol Relating to the Status of Refugees." URL https://www.unhcr.org/3b66c2aa10.

United Nations University. 2015. "Protecting Female Refugees against Sexual and Gender-based Violence in Camps." URL https://unu.edu/publications/articles/protecting-female-refugees-against-sexual-and-gender-based-violence-in-camps.html.

A Nash Household Bargaining Model

A.I The Impact on Intimate Partner Violence V

(U_m^*, U_w^*) maximizes the following expression where (U_m, U_w) is feasible:

$$\begin{aligned} &(U_m(V) - d_m)^\tau (U_w(\bar{V} - V) - d_w(\delta))^{1-\tau} \\ &s.t. \qquad (U_m, U_w) \geq (d_m, d_w) \end{aligned} \tag{3}$$

The first-order condition with respect to V entails:

$$\begin{aligned} &\tau[U_m(V) - d_m]^{\tau-1} U_m'(V)[U_w(\bar{V} - V) - d_w(\delta)]^{1-\tau} \\ &+(1-\tau)[U_m(V) - d_m]^\tau [U_w(\bar{V} - V) - d_w(\delta)]^{-\tau} U_w'(\bar{V} - V)(-1) = 0 \\ \Leftrightarrow \frac{\tau}{1-\tau}&\left[\frac{U_w(\bar{V} - V) - d_w(\delta)}{U_m(V) - d_m}\right]^{1-\tau} = \frac{U_w'(\bar{V} - V)}{U_m'(V)}\left[\frac{U_w(\bar{V} - V) - d_w(\delta)}{U_m(V) - d_m}\right]^{-\tau} \\ &\Leftrightarrow \frac{\tau}{1-\tau}\frac{U_w(\bar{V} - V) - d_w(\delta)}{U_m(V) - d_m} = \frac{U_w'(\bar{V} - V)}{U_m'(V)} \end{aligned} \tag{4}$$

A reduction in the risk of return to the origin country, δ, implies that the woman's utility if divorced, $d_w(\delta)$, increases. Thus, the ratio of marginal utilities, $\frac{U_w'(\bar{V}-V)}{U_m'(V)}$, decreases. Since U_m is increasing and concave in V, while U_w is increasing and concave in $\bar{V} - V$, the reduction in δ results in a decrease in V.

A.II The Impact on Traditional Gender Norms G

(U_m^*, U_w^*) maximizes the following expression where (U_m, U_w) is feasible:

$$\begin{aligned} &(U_m(G) - d_m)^\tau (U_w(1 - G) - d_w(\delta))^{1-\tau} \\ &s.t. \qquad (U_m, U_w) \geq (d_m, d_w) \end{aligned} \tag{5}$$

The first-order condition with respect to G entails:

$$\begin{aligned} &\tau[U_m(G) - d_m]^{\tau-1} U_m'(G)[U_w(1 - G) - d_w(\delta)]^{1-\tau} \\ &+(1-\tau)[U_m(G) - d_m]^\tau [U_w(1 - G) - d_w(\delta)]^{-\tau} U_w'(1 - G)(-1) = 0 \\ \Leftrightarrow \frac{\tau}{1-\tau}&\left[\frac{U_w(1 - G) - d_w(\delta)}{U_m(G) - d_m}\right]^{1-\tau} = \frac{U_w'(1 - G)}{U_m'(G)}\left[\frac{U_w(1 - G) - d_w(\delta)}{U_m(G) - d_m}\right]^{-\tau} \\ &\Leftrightarrow \frac{\tau}{1-\tau}\frac{U_w(1 - G) - d_w(\delta)}{U_m(G) - d_m} = \frac{U_w'(1 - G)}{U_m'(G)} \end{aligned} \tag{6}$$

A reduction in the risk of return to the origin country, δ, implies that the woman's utility if divorced, $d_w(\delta)$, increases. Thus, the ratio of marginal utilities, $\frac{U'_w(1-G)}{U'_m(G)}$, decreases. Since U_m is increasing and concave in G, while U_w is increasing and concave in $1-G$, the reduction in δ results in a decrease in G.

A.III The Impact on Divorce

If the risk of return to her origin country is decreased ($\delta' < \delta$), the woman's utility of being single increases ($d_w(\delta') > d_w(\delta)$). If the increase in her utility is so high that there is no longer a utility pair such that both spouses will prefer to remain married ($U_m > d_m(\delta')$ and $U_w > d_w(\delta')$ does not hold), then the sufficient and necessary conditions for the bargaining solution resulting in marriage are not satisfied. In this case the couple will divorce.

B Additional Tables and Graphs

Table B.I: Summary Statistics for the Population of Women Who Do Not Change Status

	All Mean	Refugees Mean	Family-Reunified to a Refugee Mean
Characteristics at Immigration			
Age	32.53	33.22	31.21
Number of Children	1.33	1.21	1.56
Days Between Application and Asylum	497.81	497.81	.
Year of Application	2002.43	2002.43	.
Year of Admission	2004.14	2004.11	2004.18
Education Surveyed			
Basic Education	0.69	0.69	0.69
Vocational Education	0.13	0.13	0.12
Academic Education	0.18	0.18	0.19
Education Not Surveyed	0.41	0.42	0.39
Origin Country			
Iraq	0.23	0.15	0.37
Afghanistan	0.14	0.12	0.18
Syria	0.07	0.08	0.05
Somalia	0.11	0.12	0.08
Origin Country Characteristics			
Female Labor Force Participation	0.29	0.32	0.24
Male Labor Force Participation	0.77	0.76	0.78
Fertility Rate	4.32	4.18	4.57
Share of Seats in Parliament Held by Women	0.12	0.12	0.13
Share of Women Married Before Age 15	0.06	0.05	0.06
Share of Women Married Before Age 18	0.25	0.24	0.27
Avg. Daily Hours of Domestic Work (Women)	4.76	4.76	4.76
Avg. Daily Hours of Domestic Work (Men)	1.56	1.58	1.48
Characteristics of Husband at Immigration			
Same Origin Country	0.78	0.73	0.83
Employment Rate	0.14	0.06	0.22
Any Employment	0.20	0.10	0.31
Labor Income (1,000 USD)	5.58	2.10	9.48
Husband's Education Surveyed			
Basic Education	0.54	0.50	0.59
Vocational Education	0.20	0.23	0.17
Academic Education	0.26	0.27	0.24
Education Not Surveyed	0.36	0.38	0.34
N	11,385	7,484	3,901

Notes: Summary statistics for the population of females who are either admitted as refugees or family-reunified to a refugee, and who do not change status. Age and number of children are measured in the year of family-reunification/asylum. Year of application and days between application and asylum is only measured for refugees. The family-reunified do not apply for asylum and thus do not have an application record. Educational attainment acquired abroad shows the distribution across different education levels for those who were surveyed about this. Origin country characteristics are measured at the country level; Labor force participation, fertility rates and the share of seats in parliament held by women are measured in the year of immigration. The share of women married before age 15/18 and average daily hours of domestic work are measured as country averages across all years available in the UN Gender Statistics and Indicators database. The characteristics of the husband are measured in the year of the woman's immigration if she is married (57 percent of the refugee women are married at immigration).

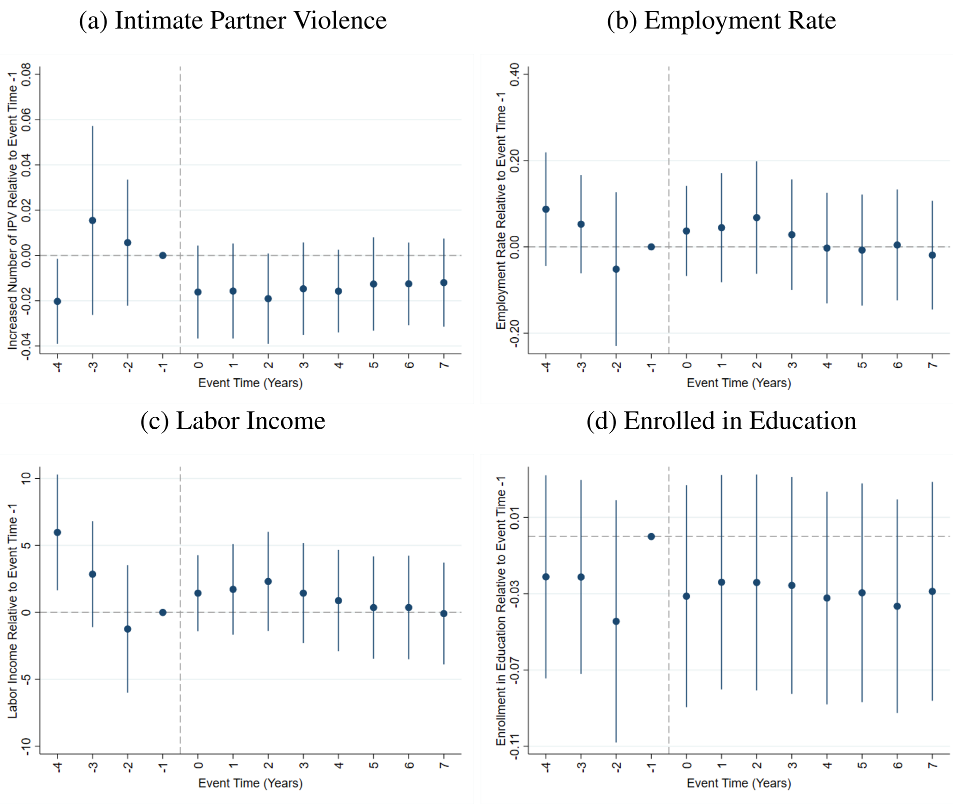


Figure B.I: Results by Marital Status, Regression Tests

Notes: The blue vertical lines represent 95 pct. confidence intervals based on robust standard errors clustered by individuals. In all panels calendar year FE and individual FE are included. The blue dots represent the estimated coefficients on the interaction term between the event time dummies and an indicator for being divorced in a given year in Panels a, b and d. In Panel c the blue dots represent the estimated coefficients on the interaction term between the event time dummies and an indicator for ever divorcing throughout the time period. This is because the risk of experiencing IPV is zero if the woman is unmarried.

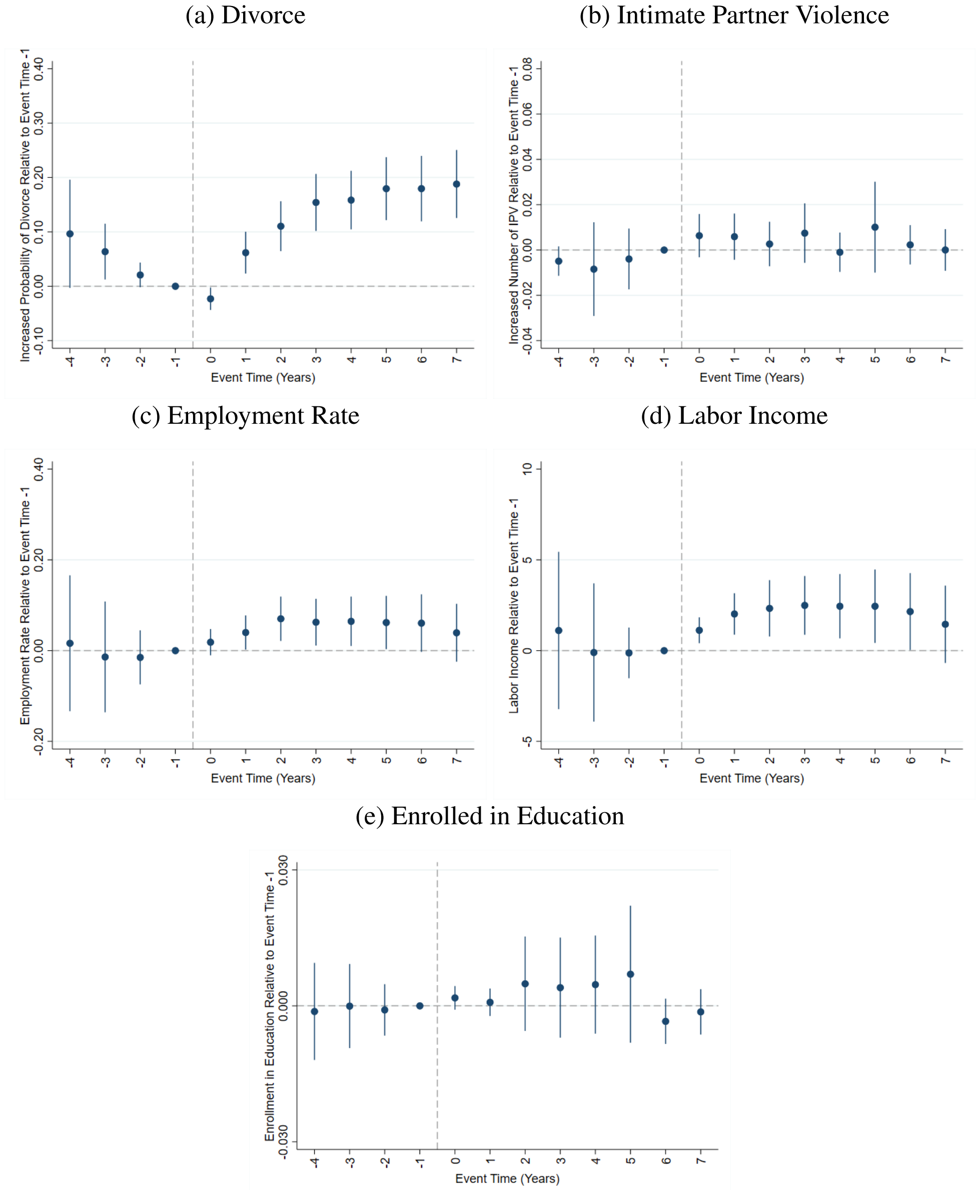


Figure B.II: Results by Female Labor Force Participation in Origin Country, Regression Tests

Notes: The blue vertical lines represent 95 pct. confidence intervals based on robust standard errors clustered by individuals. The blue dots represent the estimated coefficients on the interaction term between the event time dummies and an indicator for originating from a country with above median female labor force participation. In Panel a calendar year FE are included. In Panels b-e calendar year FE and individual FE are included. There are no individual FE in Panel a due to very little variation in the outcome variable: the majority of women only change marital status once throughout the period if they divorce.

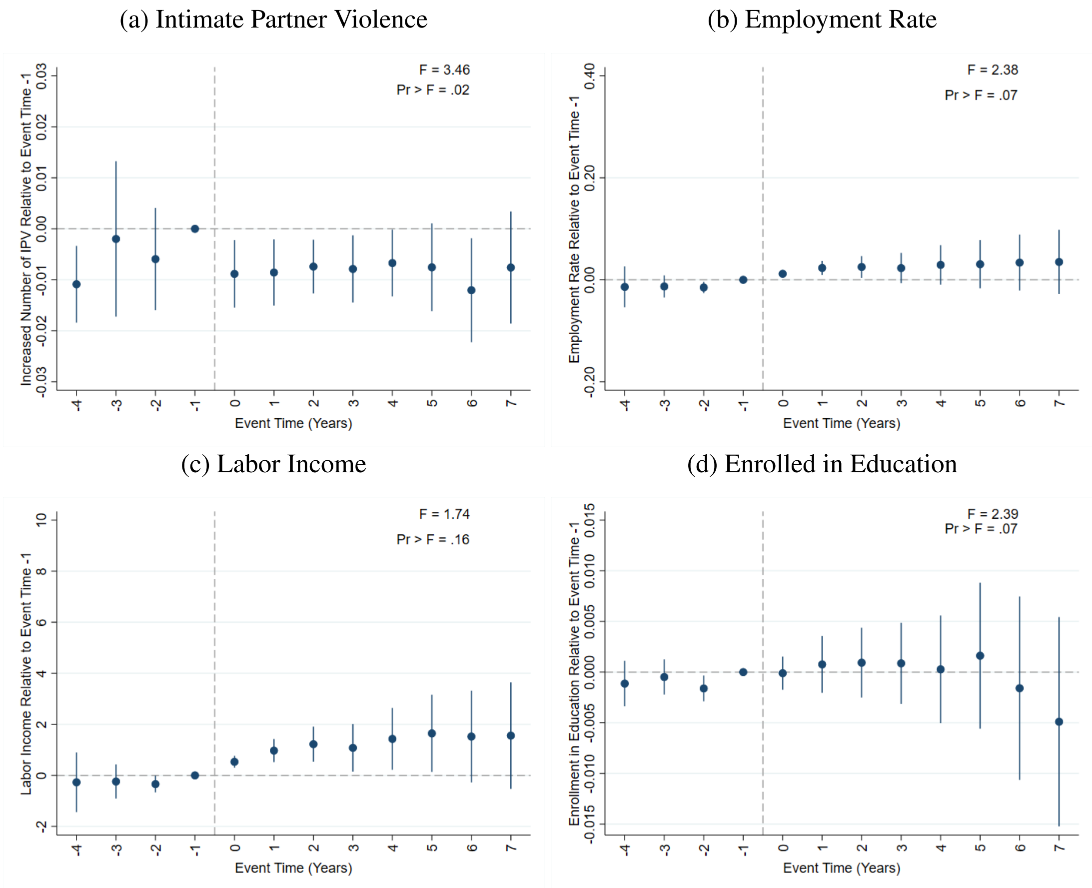


Figure B.III: Results With Calendar Year FE and Years Since Migration FE

Notes: The blue vertical lines represent 95 pct. confidence intervals based on robust standard errors clustered by individuals. In all panels calendar year FE and years since migration FE are included. The F-statistic and the p-value for the joint insignificance of the pre-event dummies are shown in the top right corner of each panel.

Chapter 4

Labor Market Rigidities and the Wage Impact of Immigration

Labor Market Rigidities and the Wage Impact of Immigration*

Mette Foged, Linea Hasager and Vasil Yasenov†

August 6, 2020

Abstract

We study the role of labor market institutions in affecting the wage impact of immigration using a cross-country meta-analysis approach. We gather information on 1,030 previously reported semi-elasticities from 54 academic studies covering 18 developed countries. We supplement this dataset with country-level institutional strength and coverage data from the OECD. Our results suggest that higher labor market rigidity mitigates the effect on relative wages of native workers with skills most similar to immigrants but exacerbates the impacts on average earnings in the economy. In other words, institutions shield native workers from distributional consequences but diminish potential benefits induced by foreign labor.

JEL Classification: D02, J08, J15, J31, J61
Keywords: immigration, wages, labor market institutions, meta-analysis

*This is a substantially revised version of the Stanford Immigration Policy Lab Working Paper No. 19-07. We thank Frida Dyred and Emma Hedvig Pind Hansen for excellent research assistance. Bernt Bratsberg, Oddbjorn Raaum, Jens Hainmueller, David Laitin, Adam Sheridan, Jan Stuhler, Birthe Larsen, Asger Moll Wingender, Morten Olsen, participants at the Third Dondena Workshop on Public Policy, seminar participants at Kraks Fond Institute for Urban Economic Research, seminar participants in the Danish Ministry of Employment, and the Immigration Policy Lab at Stanford University provided helpful comments and discussions. The Danish Ministry of Employment provided generous financial support (grant number: 6149-00024B). All errors are our own.

†Mette Foged (corresponding author): Mette.Foged@econ.ku.dk, University of Copenhagen, and IZA. Address: Øster Farimagsgade 5 Building 26, 1353 Copenhagen K, Denmark; Linea Hasager: tlh@econ.ku.dk, University of Copenhagen; Vasil Yasenov: yasenov@stanford.edu, Immigration Policy Lab, Stanford University, and IZA.

1 Introduction

Institutional structures and policies such as collective bargaining, unemployment insurance and active labor market policies play a key role in most interactions in the labor market (Nickell and Layard, 1999; Boeri and Van Ours, 2013). As such, the origins and consequences of various labor market institutions have been intense objects of analysis by economists for more than a century (e.g., Moore, 1911). More recently, economists have focused on studying the interplay between these structures and other economic phenomena. Examples include the interaction of institutions with the returns to education (Webber, 2014), income inequality (Farber et al., 2018) and wage determination (Nunziata, 2005) where researchers have documented significant synergies.

A specific aspect of the labor market where institutions may be particularly crucial, and yet are understudied, is the competition between native and foreign workers. Institutions are often meant to protect incumbent natives from competition with foreign-born labor. Specifically, policy-induced labor market imperfections and rigidities may serve as a preventive factor for wage and employment adjustments. Following a supply shock, well-designed institutions could shield native workers and provide a smooth transition back to market equilibrium, mitigating any adverse impacts caused by the foreign workers. In fact, demand for such protective institutions may endogenously arise from public pressure towards policy makers due to increased migration levels (Rodrik, 1997).

While a large body of literature analyzes the wage and employment consequences of immigration, we know little about their interaction with labor market institutions. The classic theoretical underpinnings of the literature rest on assuming perfectly competitive markets, entirely ignoring any such synergies (e.g., Borjas, 2013). This interplay is of particular relevance because real world labor markets are far from perfect and institutions may aid in absorbing foreign workers whilst minimally disrupting the labor market. The substantial returns to improved policies are highlighted by a few recent polls suggesting immigration is the top issue among election voters in Europe and the United States (e.g., European Commission, 2018; Reuters, 2018).

Two major challenges stand in the way of measuring the role of institutions on the labor market impact of foreign born workers. First, they are generally set on a national level and do not vary within countries, so a thorough analysis requires a cross-country perspective. Labor markets across the developed world vary greatly in terms of institutional scope and intensity providing a useful laboratory to study how labor market institutions affect the consequences of immigrant labor. Second, gathering comparable micro data across a wide set of countries and time periods is a virtually unattainable task.

We overcome both issues by directly relating previously estimated immigration wage effects and institutional strength in a cross-country analysis.[1] In setting up the empirical analysis, we document a strong correlation between various measures of institutional strength and, hence, think of them as all measuring the same underlying quantity – labor market rigidity. We therefore use the terms "institutional strength" and "labor market rigidity" interchangeably throughout the paper. Analyzing a suitable reform within a single country could potentially isolate the impact of one specific institution. Our cross-country approach, to the contrary, picks up the role played by clusters of strong institutions compared to more flexible labor markets. Notable examples are the Southern European models with strict employment protection and high coverage of collective agreements vis-à-vis the Anglo-Saxon pro-competitive labor markets with weaker protections and more flexible wages.

We study how labor market rigidity affects the impact of immigration on natives' wages by utilizing a cross-country meta-analytic approach. To accomplish this, we gather a novel database covering 1,030 estimates on the wage impact of foreign workers from 54 published academic articles spanning 18 countries in the developed world and several decades. We supplement this database with Organization for Economic Co-operation and Development (OECD) data on country-specific wage and employment institutions. We distinguish between "total" and "partial" wage effects because (i) they have distinct implications for the wage structure, inequality and workers' welfare more broadly (see Dustmann, Schönberg, and Stuhler, 2016) and (ii) there is a substantial difference in the two distributions of reported effects as we show in this paper. We then use a linear regression framework to relate the estimated wage effects and institutional strength while controlling for related characteristics. This type of meta-analysis is a research synthesis that attempts to extract useful conclusions from a body of diverse academic literature (see e.g., Card and Krueger, 1995; Longhi, Nijkamp, and Poot, 2005; Card, Kluve, and Weber, 2018). Importantly, we control for continent fixed effects accounting for time-invariant differences across continents and country-specific economic conditions such as GDP growth and unemployment rate. Lastly, to eliminate potential concerns about our selection of estimates and studies we control for study characteristics and use various weighting schemes as robustness checks. This is important, because – just like any meta-analysis – our finding is limited by the quality and comparability of the underlying studies.

Our main finding is that higher labor market rigidity is associated with a smaller decline in the relative wage of workers most similar to immigrants; and hence, reduces the effect of immigration on the structure of wages. At the same time, stronger institutions are negatively associated with the total

[1] Appendix B.3 contains the same analysis with employment effects. We do not find a clear pattern in how employment effects interact with institutions.

average effects on native workers, suggesting that institutional strength may hamper efficient responses to immigration. Intuitively, our findings suggest that any rigidity meant to protect incumbents will also diminish potential benefits induced by the newcomers. These may include complementarities in the production process (Ottaviano and Peri, 2012), productivity spillovers (Hunt, 2017) and skill or occupational upgrading (Peri and Sparber, 2009; Foged and Peri, 2015). The findings imply that the significant variation in institutional rigidity across countries may, at least partially, drive the substantial diversity of reported elasticities in the literature.

We advance the literature on the labor market effects of immigration by building on existing studies, confirming clear differences in the empirical distributions of partial and total effects (as proposed by Lewis and Peri, 2015; Dustmann, Schönberg, and Stuhler, 2016) and providing novel insights on the potential role of institutions in mediating the competition between native and foreign workers and hampering efficient responses by workers and firms. In doing so, this paper presents a comprehensive analysis of the synergies between institutional rigidities and the consequences of immigrants, filling in an important gap in an otherwise large and mature body of literature. Consistent with the previous evidence (Angrist and Kugler, 2003), we find that labor market rigidity exacerbates the effect of immigrants on the average wage earner. Extending this analysis, we are the first to show that institutions are helpful in mitigating immigration-induced wage changes among various native skill types. Although several important recent studies have strayed away from the classic perfect competition framework and begun analyzing the role of individual institutions, a thorough analysis has so far been elusive.

In a series of influential meta-analyses on the labor market effects of immigration, Longhi, Nijkamp, and Poot (2010a; 2008b; 2008a; 2005) show that the estimated elasticities in the United States tend to be smaller in absolute value than in Europe. The authors speculate this disparity may stem from differences in labor mobility and/or labor market institutions but do not test this hypothesis. Perhaps most similar to our study is the paper by Angrist and Kugler (2003) who use the Bosnian wars as quasi-experiments to analyze the interaction of institutions and employment losses induced by labor supply shocks. Reduced market flexibility (as measured by higher firing costs, replacement rates and barriers to entry) is associated with increased employment losses while wage consequences are not analyzed. Next, Brücker et al. (2014) use a highly structural wage-setting approach to study the immigration earnings effects in three European countries - Denmark, Germany and the United Kingdom. They find a significant role for wage flexibility in determining its magnitude, but do not identify the effect of particular institutions or policies. Lastly, Edo and Rapoport (2019) document an interesting interaction whereby a lower minimum wage exacerbates the immigration-induced earnings and employment consequences in the United

States.[2]

A common theme among these studies is the significant role of institutional structure in altering the competition forces between incumbent workers and newcomers. Our paper builds on and extends this body of work, distinguishing between the total average and relative partial wage effects of immigration, analyzing a combination of wage and employment institutions and employing a strictly data-driven approach. Consistent with these earlier studies, we show that reduced labor market flexibility is associated with higher average wage losses due to immigration. In addition and novel to the literature, we go one step further to document that labor market rigidities mitigate impacts on relative wages in the labor market. For recent reviews of the broader literature of the labor market consequences of immigration on domestic economies see NAS (2017), Dustmann, Schönberg, and Stuhler (2016), and Lewis and Peri (2015).

We continue in section 2 where we outline a general version of the theoretical models in the literature on the labor market impact of immigration. We also define and describe in detail the total and partial effects on wages which we analyze later on. We continue with section 3 where we present the labor market rigidity variables from the OECD data. Next, in section 4 we describe the database of estimates on the wage impact of immigration. Section 5 lays out the equation we estimate. Finally, section 6 presents our results and concludes the paper.

2 Theoretical Background

The literature on the labor market impacts of immigration builds on classic factor demand theory. It is generally assumed that capital has an equal degree of substitutability with all workers.[3] Hence, without loss of generality, we can write the production function in each region r (a national or a local economy) as:

$$Y_r = G(A_r, K_r, L_r), \tag{1}$$

$$L_r = F(A_{r1}, L_{r1}, A_{r2}, L_{r2}, ..., A_{rs}, L_{rn}), \tag{2}$$

[2] A strand of literature related to our study analyzes the extent to which these institutions affect the labor market outcomes of immigrants (e.g., Kogan, 2006; Bergh, 2017; Fleischmann and Dronkers, 2010).

[3] See Lewis (2011, 2013); Lafortune, Lewis, and Tassada (2019) for analyses of relative complementarity between capital and skilled (college-educated) workers.

where Y_r is output, K_r is physical capital, L_r is a labor aggregate and A_r captures total factor productivity (TFP) as well as the relative productivity of capital.[4] The aggregate labor factor combines n distinct labor inputs $L_{r1}, L_{r2}, ..., L_{rn}$ and their relative productivity parameters $A_{r1}, A_{r2}, ..., A_{rn}$. Depending on the study, workers may be differentiated by education or occupation and maybe experience (see Appendix A.1 for details) and we refer to each type as a skill cell s. Immigrant (IMM) and native (NAT) workers can be correctly assigned to these cells and are perfect substitutes within them: $L_{rs} = L_{rs}^{NAT} + L_{rs}^{IMM}$.[5] Both functions $G(\cdot)$ and $F(\cdot)$ are homogeneous of degree one (i.e., exhibit constant returns to scale), strictly increasing and strictly concave (implying some degree of substitutability between inputs). Furthermore, native labor supply is fixed such that the only source of variation in L_{rs} is the immigrant inflow that is assumed to be exogenous.

Profit maximization and perfect competition imply that workers are paid their marginal product. Ignoring the subscript r, we have:

$$w_s = \frac{\partial Y}{\partial L}\frac{\partial L}{\partial L_s} = G_L F_s. \tag{3}$$

Wages in each skill cell s are determined by the marginal productivity of the labor aggregate (G_L) and the marginal contribution of skill group s to that labor aggregate (F_s). The first term can change if TFP changes ($dA \neq 0$) or the capital-labor ratio changes ($d\frac{K}{L} \neq 0$). The second term can change if the relative skill supply is altered ($\frac{dL}{L} \neq \frac{dL_s}{L_s}$).

We distinguish between two distinct effects of immigration on wages. The "total effect" on natives in skill cell s includes all changes occurring on the right-hand-side of equation (3) in response to immigration. It captures the change in natives' wages due to immigration when allowing all features of the economy to respond to the arrival of foreign-born workers. Hence, the total effect accommodates changes in both G_L and F_s. The "partial effect," on the contrary, isolates the direct impact of an immigration-induced supply change in the same skill cell, holding fixed all cross-cell effects (in F_s) as well as impacts on TFP and the productivity of capital (in G_L). It thus separates the direct impact on F_s from all other effects.[6] We will now examine each type of effect in turn and how they are estimated.

[4]The simplest model features identical workers. In such a world, by construction, there are no distributional effects of immigration and capital flexibility pins down the impact on wages. Following an influx of foreign-born workers, wages decrease in the short run due to declining marginal product of labor. In the long run, capital adjusts and there is no wage effect. This is clearly too simplistic. Hence, we proceed here with n distinct types of workers.

[5]Dustmann, Schönberg, and Stuhler (2016) discuss the implications of downgrading of immigrants' skills.

[6]Note that both the "total" and "partial" wage effects have analogous employment counterparts when native employment is elastic. In this case, we need to know the slope of the labor demand curve (the total derivative of equation 3) as well as the slope of the labor supply curve in order to calculate the wage and the employment effects of immigration (see e.g., Dustmann, Schönberg, and Stuhler, 2016).

The direct partial effect is always negative, unless the assumption of perfect substitutability between immigrants and natives within cells does not hold.[7] Studies utilizing the so-called *national skill-cell approach* (e.g., Borjas, 2003) use variation in the immigrant share across skills cells and time in national labor markets to obtain such a partial effect. Specifically, they isolate this theoretical parameter by controlling for (i) skill-cell fixed effects (capturing level-differences in wages across skill cells), (ii) aggregate time fixed effects (adsorbing TFP growth and changes in the relative productivity of capital), and (iii) worker-category-by-time fixed effects (capturing impacts of immigration shared within broader categories of workers).[8] By comparing natives' wages in cells more versus less exposed to immigration, the estimated parameter is informative about changes in *relative* earnings and we will sometimes refer to it as the relative wage effect. This is important because complementary workers in other (indirectly affected) cells will magnify the estimated parameter compared to the direct partial effect.

The total effect is more complex to analyze. In general, native workers with most similar skills to those of immigrants experience a wage loss, while the rest may gain or lose depending on the changes in relative supplies and the elasticities of substitution across worker types. Empirical papers rarely isolate total wage effect estimates by detailed skill cells. Instead, they estimate an *average* total effect.[9] The sign on this average effect is theoretically ambiguous and it could even be positive. For instance, an influx of low-skilled immigrants may increase natives' average wages if low-skilled native labor is sufficiently scarce.[10] Studies regressing natives' wages on immigrant shares across regions in a difference-in-differences design or a panel data setting follow the so-called *spatial approach*.[11] This approach was pioneered by studies such as Card (1990); Altonji and Card (1991) and later followed by several others including Foged and Peri (2015) and Peri and Yasenov (2019). These studies capture the total effect of immigration in regions experiencing a surge in foreign-born compared to localities with lower immigration. Acknowledging the workforce heterogeneity, we will sometimes refer to this as an average total effect or simply an average effect.

Lastly, a third category of studies utilizes variation of both skill cells and regional labor markets (the so-called *mixture approach* e.g., Borjas, 2006; Card, 2009; Glitz, 2012) to isolate the labor market

[7]Ottaviano and Peri (2012) provide evidence of imperfect substitutability in the classical education-experience cells formulation. This shifts the negative (i.e., competition) effect to be experienced primarily by earlier immigrants rather than natives. Moreover, Peri and Sparber (2009) and Foged and Peri (2015) explain this by different occupational specialization of immigrants and natives.

[8]See Appendix A.1 for a detailed description of specifications included as capturing a relative partial effect of immigration.

[9]Notice that $w = \frac{\partial Y}{\partial L} = G_L$ is not the average wage but rather an artificial theoretical concept because L is an efficiency-weighted aggregate of n distinct labor inputs (not body counts).

[10]See Appendix A.2 for the additional assumptions required and a formal derivation of this result.

[11]TFP, modes of production and factor supplies in equation (1) and (2) and hence wages in equations (3) differ across regions, highlighting the importance of region fixed effects in the spatial approach, see Appendix A for a detailed discussion of the estimates we categorize as total effects.

impacts of foreign-born workers. For our purpose it is sufficient to note that these specifications also capture a partial effect since impacts shared within broad categories of workers are again absorbed by fixed effects.

There is considerable agreement in the literature on the importance of distinguishing between total and partial labor market effects of immigration. See, for instance, recent reviews by (Peri, 2014; Lewis and Peri, 2015; Dustmann, Schönberg, and Stuhler, 2016). We supplement these papers as well as earlier meta-analyses (Longhi, Nijkamp, and Poot, 2010b; 2010a; 2008a; 2008b; 2005) by summarizing a large number of estimates categorized into partial and total effects. Specifically, we show that, in addition to the theoretical distinction, their empirical distributions are rather different as well (see Section 4.4).

The canonical labor demand model, outlined above, is sometimes augmented with various adjustment mechanisms. In more elaborate settings, a number of additional responses by workers and firms may mitigate and even reverse the standard predictions. For instance, endogenous technological choice (firms choosing modes of production, e.g. the A_{rs}'s) and trade can both potentially offset immigrants' effect on wages (Card and Lewis, 2007; Lewis, 2013; Dustmann and Glitz, 2015). Additionally, immigration induced innovation and knowledge spillovers positively affecting TFP as well as efficiency gains from specialization can potentially lift wages (Ottaviano and Peri, 2006; Peri and Sparber, 2009; Peri, 2012; Lewis, 2013).[12] Very little empirical evidence exists to gauge the importance of these margins of responses to immigration. In most countries from our sample, however, immigrants are less skilled than natives. Therefore, trade, choice of technique, complementarity and task-specialization seem more relevant than knowledge spillovers, which is usually attributed to high-skilled immigration. Lastly, real world labor markets are far from perfectly competitive and institutional rigidities likely interact with immigration. We now turn to a discussion about the way we measure these rigidities in the empirical analysis to follow.

3 Labor Market Rigidity Measures

There are two main types of institutions in the labor market – ones that regulate prices (i.e., wages) and ones that act on quantities (i.e., employment). In practice they coexist and there is considerable variation in their strength and scope across countries. The OECD has developed detailed synthetic measures of wage and employment protection (OECD, 2018). This is the primary source of information for labor market rigidities we use in the empirical analysis. Table 1 lists the countries in our database of wage

[12] Ottaviano and Peri (2006) is an example of a paper that does not meet the inclusion criteria explained in section 4 because immigration is measured with a diversity index not as the proportion of immigrants in the population.

impacts of immigration (rows) and the distinct institution variables we use to measure labor market rigidities (columns). The first three columns present measures of employment rigidities while the last two display wage-protecting policies. All values reflect country averages for the period 1998-2016.

Overall, Table 1 shows there is substantial variation in institutional strength across European countries, while the United States stands out as having a relatively pro-competitive and flexible labor market. The Employment Protection Legislation (EPL) indices are measured on a scale from 0 to 6 and capture respectively protection of workers on regular/permanent contracts against individual dismissals (column 1) and the strictness of requirements for collective dismissals (column 2). Higher values correspond to stronger protection. Southern European countries such as Spain, Portugal and Italy have particularly strong employment protection legislation. Workers on regular contracts are by far the most protected against individual dismissals in Portugal (4.52) and the definitions and procedures regarding collective dismissals are particularly strict in Italy (4.10) and Spain (3.73). Canada, the United States, and the United Kingdom, on the other hand, are the least regulated by these measures. Countries with stronger EPL have lower job tenure as measured by the average number of years workers have been in their current or main job/employer (column 3).[13] Note that longer job tenure corresponds to lower employment mobility (higher rigidity). The mean workplace tenure in our sample is roughly 10 years. The longest average stay is found among Southern European countries (e.g., 12.22 in Portugal and 12.03 in Italy), while in the United States it is only 7.51 years.

The collective bargaining coverage rate (column 4) is defined as the share of workers covered by collective agreements among all workers with the right to bargain. It is a direct measure of wage rigidity commonly used in the literature (e.g., Avouyi-Dovi, Fougère, and Gautier, 2013), affecting a wider spectrum of the labor market than other price institutions such as the minimum wage. An alternative and related measure of wage rigidity is trade union membership. We prefer collective bargaining coverage because it captures the actual share of the workforce covered by collective agreements. Trade union membership is misleadingly low for some countries with high wage rigidity and, unlike collective bargaining coverage, it has decreased over time due to increased use of extension clauses. For example, 88 percent of workers in France are covered by collective agreements but less than 10 percent are members of a union. Table 1 shows that European workers are covered by collective agreements at much higher rates (e.g., 97% in Austria and 88% in France) than workers in the United States and Canada (21% and 33% respectively) while the sample average is 56%. Finally, the net replacement rate (column 5) is the

[13] The data source for the United States is the Job Tenure Supplement of the Current Population Survey and the sample used to calculate it is all people aged 16-64 in the labor force.

guaranteed compensation rate for a typical two-earner household (as percent of their salaries) in the initial phase of unemployment. Countries with more generous compensation programs likely have higher wage rigidity since unemployment benefits act as a wage floor. Compensation levels are higher in Continental Europe (e.g., 93% in Denmark and 92% in Germany) and lower in the Anglo-Saxon countries, most notably the United Kingdom (65%) and Australia (65%).

Next, Table 2 shows the correlation matrix of our five measures of institutional strength which we also refer to as labor market rigidity. The correlation coefficients are always positive and usually in the 0.2–0.8 range. The quantity and the price institutions are strongly correlated across countries. For instance, protection of regular workers against individual dismissals is highly correlated with collective agreements coverage (0.68) and they are both strongly associated with average job tenure (0.73 and 0.59 respectively). Hence, we think of each institution as a measure of the same underlying concept – labor market rigidity – viewed from a slightly different angle. Our goal in the empirical analysis is to correlate the wage impact of immigration and broad labor market rigidity, rather than focusing on individual institutions and policies. This is important since taken together wage and employment institutions have different implications than each of them in isolation.

The number of jobs in the economy is ultimately fixed and immigration simply creates unemployment if wages cannot adjust. Preserving pre-immigration employment would instead imply that wages adjust. Together EPL and rigid wages can instead protect incumbent workers and reduce the volatility of wages and employment. Understanding how different types of workers are affected in the interplay between immigration and institutional rigidity is a challenge. It depends in complicated ways on how these institutions privilege subsets of workers and how immigration changes their relative supplies. The impacts on the structure of wages (the relative partial effects) could be smaller in clusters of strong employment protection and high coverage of collective agreements as suggested in the empirical analysis below. Intuitively, such institutional rigidity limits the ability of the economy to respond to immigration and therefore also reduces potential benefits (captured in the total effects).

4 Estimates of the Wage Impact of Immigration

4.1 Sample

Building a database of estimates on the labor market effects of immigration requires selecting study inclusion criteria based on a few key priorities. With the goal of studying labor market rigidity, we focused on sampling studies from a wide variety of labor market settings across the developed world and from recent decades where the institutional data is available (not historical immigration episodes). Moreover, to ensure study (and therefore data) quality, we relied on peer-reviewed publications rather than unpublished manuscripts. We refer the reader to the Online Data Appendix for a detailed description of our database including criteria for the measurement of immigration and the way collected other study characteristics while Appendix Table B.1 provides a list of the included studies. Once the potential candidates were collected we determined for each empirical specification whether reported effect sizes could be classified as partial effects, total effects or none of those (exclusion).

Our database consists of 1,030 estimated wage effects of immigration from 54 studies published between 1990 and 2020 spanning 18 European countries, the United States, Canada, Turkey, Israel and Australia.[14] When available, we recorded data on the corresponding standard error, t-statistic (associated with the null hypothesis of the true effect equals zero), underlying sample (entire workforce or by education level), level of variation in the dependent variable of interest, publication outlet, estimation method (OLS, IV, Difference-in-Differences or First-Differences) and empirical strategy (natural experiment, Bartik type instruments, other instruments and OLS). To gauge study quality, we also included journal impact factors measured by IDEAS/RePEc (2020), a widely used index. We also gathered data on 432 estimates of the employment impacts of immigration from 29 studies covering 9 countries which we describe in the Appendix B.3.[15]

Differences in the independent variable of interest across studies require that a few adjustments be made to ensure the reported effect sizes are measured in comparable units. First, included studies either publish elasticities or semi-elasticities. The distinction is subtle but important. The former measure the percentage change in natives' wages induced by a percent increase in immigration while the latter do the same for a percentage point increase in foreign-born workers. Following Longhi, Nijkamp, and Poot (2005), we convert all elasticities to semi-elasticities since this is the more commonly used metric in the literature. Second, the wage effects are sometimes reported for growth rates rather than levels (of immigration). In that case, we rescale the parameter to reflect a one percentage point increase in the proportion of foreign-born. Hence, our estimates are interpreted as the percentage change in wages (or percentage point change in the employment rate) for a one percentage point increase in immigrants as a share of the labor force. The Online Data Appendix contains a complete description of the operations

[14]The majority of the collected wage effects are from a regression with either hourly, daily or weekly earnings as the dependent variable (332 out of 613 partial and 381 out of 417 total wage effects). The remaining estimates are based on monthly, quarterly or yearly earnings and for three papers (34 estimates) it was not possible to obtain a clear definition of the wage variable from the primary study.

[15]These numbers combine the impacts on the employment rate e with the negative of the impacts on the unemployment rate u, assuming the participation rate p is close to one and unaffected by immigration (by definition $e = p(1 - u)$.

and technical details involved in creating comparable estimates across studies.

While we attempt to do our best effort in assuring comparability across studies, an inherent limitation of the meta-analysis approach is the inevitable differences across samples and regression specifications for which we cannot account. For instance, included control variables or fixed effects across studies are rarely exactly identical.[16] The sample definitions vary as well – e.g., 16-65 versus 21-54 years old, working age population versus labor force, both gender versus only men etc. We have no reason to believe these remaining differences are systematically associated with labor market rigidity and hence biasing our results in a certain direction.

4.2 Weighting

The sheer differences across studies raise the question of how much importance to place on each effect size. For instance, should the semi-elasticities from a paper reporting only five estimates receive equal weights to ones derived from a paper with 40? Similarly, should articles published in the top academic journals, which have undergone a more rigorous review process and are supposedly of higher quality, be weighted similarly to ones from lower ranked outlets? Lastly, one can argue that more precisely estimated semi-elasticities should also be given higher importance as they contain less statistical noise.

To verify the robustness of our results and to account for discrepancies across studies we introduce several different weighting schemes. First, we assign each effect size an equal weight and we refer to these as our "unweighted" results. Second, we give equal weights to each article, and we call these our "studies" weights. This procedure uses the number of reported semi-elasticities per article as an inverse probability weight and it, intuitively, down-weighs effect sizes coming from studies reporting many such semi-elasticities.[17] Third, we account for study quality by using journals' impact factors data from RePEc and utilizing the inverse score as a weight.[18] We refer to these as "impact" weights. This strategy assigns higher importance to estimates published in higher ranked journals. Fourth, we adjust for the precision of the included estimates by using the inverse standard error as weighting factor. These are our "precision" weights. Finally, we construct a "combined" weight in which we multiply together the preceding three weighting factors. This measure gives higher importance to effect sizes (i) published in higher ranked outlets, (ii) which are more precisely estimated and (iii) come from papers with fewer published semi-elasticities. As such, it combines the virtues of all specifications. We conduct

[16]However, we lay out some minimal requirements in Appendix A.1 that the included effect sizes need to meet in order to be included as either partial or total effects.

[17]See the Appendix Table B.1 for a list of included studies and the associated number of estimates obtained from the paper.

[18]Note that higher scores correspond to lower journal ranking.

our analyses separately with each of these weighting factors.

4.3 Publication Bias

A key concern for every meta-analysis study is the extent to which the sample of estimates is representative of the entire population. It is well-known that, all else equal, studies with statistically significant estimates are more likely to be published in academic outlets. Unfortunately, this introduces an inherent bias towards more significant estimates even in a random sample of published studies.

To analyze whether such bias may be present in our sample, we follow the standard practice in the literature (e.g., Stanley and Doucouliagos, 2014; ; Stanley and Doucouliagos, 2012; Christensen and Miguel, 2018) and conduct two tests. First, Figure 1 presents a funnel plot of our estimates on the wage impact of immigration. This is a scatter plot of estimated precision (inverse standard error) against effect magnitude. For clarity, we have removed a few outliers with precision values greater than 80. Estimates generated from smaller, noisier samples form a more dispersed base while more precisely estimated effects are more narrowly clustered around the "true" effect. In absence of publication bias, the scatter plot should be symmetric around this value, the median represented by a red dashed vertical line. For instance, if negative estimates were more likely to be published than positive ones, we would expect an excess mass on the left of this vertical line. While several estimates stand out further away from the red line, the scatter plot seems rather symmetric, indicating no severe publication bias.

Second, to formally test the symmetry of the funnel plot we proceed with a regression analysis. Following the meta analysis literature in economics (e.g., Stanley and Doucouliagos, 2012; Christensen and Miguel, 2018), we regress the t-statistic on a constant and the standard error. The coefficient of interest is that on the constant, which quantifies the correlation between precision and effect size. In a setting without publication bias, it would not be statistically significantly different from zero. Table 3 presents the results for the different weighting schemes we use throughout the paper and explain above. Standard errors are clustered by study and shown in parentheses. Only one of the five coefficients is marginally statistically significant (at the 10% level). The rest are not significant, confirming the visual inspection concluding symmetry of the (unweighted) funnel plot.

Overall, the results presented in this subsection assert that publication bias is likely not a major concern for the analyses we conduct. For the purpose of our study, publication bias, even if present, would only be problematic if it is correlated with labor market rigidity across countries. We have no reason to believe this would be the case. Lastly, Figure B.1 and Table B.4 in the Appendix present the same results for our database on the employment effect of immigration where the evidence for publication bias is slightly stronger.

4.4 Summary Statistics

Figure 2 visualizes the distribution of the estimated wage effects of immigration. Each panel weights the semi-elasticities with a different factor denoted in the figure title. A few key facts are worth highlighting. First, the impact of immigration is centered around zero in all specifications. Second, the distribution of partial effects is always shifted to the left compared to the one of total effects. Furthermore, the medians and the means of the partial wage effects are always negative, while those of the total impacts are smaller in magnitude and usually positive. Third, down-weighing noisy estimates and assigning larger weights to papers in better ranked journals result in eliminating outliers and narrowing the distributions. Lastly, the distributions of partial and total employment effects (shown in Appendix Figure B.2) look overall similar. However, all measures of centrality are negative and small, and there is relatively more mass around zero (i.e., fewer outliers).

Table 4 presents additional descriptive statistics for partial (Panel A) and average (Panel B) wage impact estimates.[19] Each column shows a different statistic denoted in the header. The top row displays information on our entire sample of 1,030 estimates, divided into 613 partial and 417 total wage impacts which were described in the histograms. A one percentage point increase in the share of foreign workers, on average, lowers the wage of unskilled workers by 0.38 percent relative to the wage of skilled workers, ceteris paribus (Figure 2 Panel A).[20] The same influx is associated with a 0.14 percent increase in the average labor market wages (Figure 2 Panel B). In other words, foreign workers assert stronger pressure on the relative earnings between skill groups than on the average native. This is an interesting finding which has not been thoroughly documented in previous summaries of the literature. It is indeed the prediction of the canonical heterogeneous labor model (section 2) in which immigration-induced changes in relative skill supplies change the relative earnings differential while the overall effect may be close to null or even positive. This result further motivates our distinct treatment of the two types of semi-elasticities in the regression analysis. The pooled mean (median) in our sample (not shown here) is -0.15 (-0.03) which is very close to the -0.12 (-0.04) reported in Longhi, Nijkamp, and Poot (2005) based on a smaller sample of 345 effect sizes from 18 papers.

The next three rows of Table 4 break the collected effect sizes down by continents. The partial

[19] Appendix Table B.5 does the same for employment effects.

[20] This interpretation is relevant for the majority of the countries in our sample because their immigrant population is less skilled than their native born population. More generally, the parameter captures the effect on the workers more versus less similar (directly exposed) to immigrants.

(average) wage effects are more negative (positive) in North America (means equal to -0.58 (0.16)) relative to Europe (-0.26 (0.13)) or the rest of the world (-0.03 (0.11)). Most of our estimates come from Europe (564) and the United States and Canada (403). Moreover, the relative partial wage effects are close to zero (median 0.02, mean -0.05, N=74) in Southern Europe which we describe in section 3 as having particular strong institutions and large in the more flexible Anglo-Saxon economies (-.38, -.51, 348). At the same, average total wage gains are positive but small in Anglo-Saxon economies (.17, .17, 223) and negative Southern Europe (-.49, -.74 22). This provides descriptive suggestive evidence of the idea that labor market rigidity reduces impacts on the wage structure and reduces potential gains associated with immigration.

We continue with a break down by empirical strategy. We distinguish natural experiments studies from other analyses and categorize IV into those based on natural experiments, Bartik type and other. The residual (called OLS in the Table) is not IV and not based on any natural experiment. Partial effects coming from natural experiments or using Bartik style instruments are smaller in magnitude than those obtained with other instrumental variable strategies while OLS effects tend to be within the same range. OLS and Other IV seem to be positively biased in spatial studies compared to the Bartik type IV and natural experiments. This is consistent with the idea that booming labor markets attract immigrants. Cross-tabulations with journal scores and impact factors (not shown here), reveal that effect sizes based on a natural experiment in an IV estimation are by far the best published followed by Bartik type IV estimates. Hence, one may think of the impact weight as giving most weight to careful natural experiment studies.

The next sub-panel summarizes the estimates by effects on high and low skilled natives as defined in the underlying studies. Unsurprisingly, low skilled natives are experiencing stronger competition with immigrants (means equal to -0.81 and 0.06) than high skilled workers (0.16 and 0.14). Finally, we see notable differences in reported wage effects of immigration by academic journal quality. Namely, papers published in Top 50 outlets feature more negative semi-elasticities than the rest. Table B.5 in the Appendix presents analogous summary statistics for our database of employment impacts of immigration.

5 Empirical Strategy

We estimate the following equation:

$$y_{ict} = \alpha + \beta \cdot institution_c + X'_{ct}\gamma + Z'_i\psi + \lambda_r + \varepsilon_{ict},$$

where the subscripts i, c and t refer to effect size, country and time period, respectively. The outcome variable y_{ict} is the estimated immigration wage effect and $institution_{ct}$ is an institutional variable of interest described in section 3. Next, the vector X_{ct} includes GDP growth and unemployment to control for national time varying economic conditions affecting labor market tightness and institutions. These variables account, for instance, for situations in which economic booms absorb foreign workers more smoothly and spur lower institutional strength, such as lower coverage of collective agreements. The vector Z_i contains study/semi-elasticity characteristics such as data frequency, estimation method, and the skills of natives. The term λ_r is a vector of continent dummies for Europe and the rest of the world, excluding North America as the reference group. Lastly, ε_{ict} is an idiosyncratic mean-zero error term.

The coefficient of interest is β. We estimate this equation with weighted OLS using the various weighting factors described in section 4.2 separately for partial and total wage effects. The standard errors are clustered by studies but do not account for the fact that our outcome variable is already estimated (as in Card, Kluve, and Weber, 2018). In an ideal scenario, we would like to estimate the interaction between the wage returns to immigration and institutional strength within countries. Nevertheless, given that institutions are set at the country level, this relationship identifies β from cross-country variation within continents (while controlling for national economic conditions, X_{ct}). This is perhaps a limitation but we believe it is the closest we can get to an association without wiping out the useful part of the variation in the data.

6 Results

6.1 Main Results

Table 5 displays our main results. Panel A shows the impacts of labor market rigidity on the partial wage effect of immigration and Panel B on the total wage effect. Each entry is an estimated coefficient $\hat{\beta}$ of an institution variable denoted in the left column. Each column corresponds to a different weighting scheme denoted in the column header. The rows below each regression coefficient denote the sample size (N) and the included controls. Standard errors are shown in parenthesis and are clustered by paper. All regressions control for study characteristics, region fixed effects and local economic conditions. To mitigate the influence of a few large outliers, we have removed semi-elasticities larger than 3 in absolute value. We present robustness checks varying these specification choices.

Nearly all coefficients in Panel A are positive and some of them are statistically significant. This suggests that institutional rigidity is effective in mitigating the relative wage changes induced by foreign workers. The first two sets of results show the regression coefficients of our employment protection legislation variables for individual and collective dismissals, respectively. A one unit increase in this index – a change equivalent to more than half the difference between the United States and Europe – is associated with a 0.03-0.4 percentage point increase in the partial wage effects of immigration. Four estimates for collective dismissals are statistically significant while the rest are mostly positive but not distinguishable from zero. The third set of estimates displays the effects of average job tenure in years which proxies for overall employment rigidity. An increase of this variable by one year – around a tenth of the sample mean – is associated with 0.22-0.32 percentage points higher relative wage effects of foreign-born workers and all but one coefficients are statistically significant.

Next, the forth row presents the results for wage rigidity as measured by collective bargaining coverage. The coefficients are all positive but noisily estimated and not significant. A one percentage point increase in the share of workers covered by a collective bargaining agreement leads to a weak increase in the relative wage effects of immigration by 0.01-0.02 percentage points. Note that such a change in coverage is minor since the sample mean of collective bargaining coverage is 56 percent, and even within Europe we see differences of about 50 percentage points. The last row shows the same results for the net replacement rate as percent of wages. This is the only variable in this panel that does not show a clear pattern. Two of the coefficients are positive and the rest are negative while none is statistically significant.

Panel B tells the opposite story – our preferred estimates (last column) suggest that labor market rigidity is associated with lower total wage effects of immigration. This result is consistent with Angrist and Kugler (2003) who find that stronger institutions worsen the employment effects of immigration in European countries ("spatial approach"). The association presented here is stronger than the relationship shown in Panel A. The first two rows in Panel B shows a unit increase in employment rigidity indices is associated with 0.02 - 0.78 percentage points lower wage effects of foreign workers. Regardless of the weighting factor, all signs are negative, pointing to the same conclusion. Similarly, the third row shows higher employment rigidity (longer average job tenure) leads to worse average wage consequences of immigration. An increase in the average tenure by one year is associated with a lowered semi-elasticity of about 0.09-0.41. All five coefficients are statistically significant. Next, our first wage rigidity variable – collective bargaining coverage – is the only one that does not show a strong pattern in this panel. The coefficients change sign depending on the weighting factor and are not statistically significant. Lastly, we find evidence that the higher net replacement rates are associated again with lowered total wage effects of immigration. A single unit increase in this variable is related to a 0.5-0.7 decrease in the associated semi-elasticity.

Overall, Table 5 suggests that higher labor market rigidity exacerbates the total wage impacts of immigration while shielding from redistribution consequences. In other words, labor market rigidity protects incumbent native workers but also dampens the potential benefits induced by the immigration. For some of the variables, this pattern is stronger and holds across various different weighting factors.

6.2 Robustness Checks

We conducted several robustness checks which we present in Tables B.2 and B.3 in the Appendix. All columns use the combined weights. First, we ran the same analysis without excluding outlier semi-elasticities larger than 3 in absolute value. The results are shown in Table B.2 and are largely similar. The broad patterns we observe hold but the coefficients are generally more noisily estimated as expected with the presence of large outliers. Next, Table B.3 presents a battery of additional robustness checks. In column 1 (in both panels) we replicate the results from Table 5 described above. In columns 2 and 3 we change the sample – by excluding estimates using variation across industries (and skill cells) and based on OLS regressions, respectively. Next, in columns 4 through 6 we change the controls by excluding in turn study characteristics, proxies for national economic conditions and continent dummies. While there is some idiosyncratic variation in the estimates across columns and panels, they rarely flip signs and the general patter broadly holds. This is especially true for the results in Panel B where most coefficients are negative and statistically significant.

6.3 Discussion

We measure labor market rigidity as brought about by more pervasive institutions, regulations and policies and investigate their role in intermediating the competition forces between native and foreign workers. We find that strong labor market institutions are associated with mitigating relative wage effects caused by immigrants but mildly exacerbate effects on average wages. In other words, these institutions seems effective in shielding domestic economies from redistributional consequences and have modest adverse impacts for the average worker. Hence, we document an important role of institutions in explaining differences in labor market impacts of immigration across countries.

Importantly, our study suffers from several limitations. First, we only consider the wage effect of immigration, thus painting only a partial picture of the role of institutions in the labor market impact of foreign workers. It is possible that quantitatively important adjustments take place through changes in employment, and these might well interact with labor market rigidity. Yet, our paper is silent on their role. Second, the limited statistical power of our analysis does not allow us to make concrete statements about the magnitudes of the effects we observe. Future research should pin down which type of institutional rigidity plays the largest roles in the pattern we document. Lastly, just like any other meta analysis, we are not able to fully control for idiosyncratic study-specific characteristics. To the extent that they may be biasing our results, we note that they have to be correlated with country level labor market rigidity. Moreover, the results from the various robustness checks we present in the appendix broadly follow the patterns we observe in the main specification suggesting these study differences may not be quantitatively important. With these limitations notwithstanding, we believe the general pattern we find is informative and novel.

Our main result entails three broad implications. First, labor market institutions can be adequate and useful tools to policy makers concerned with the redistributional ramifications of globalization and immigration in particular. Second, a thorough understanding of the interaction between labor market rigidity and wage effects is fundamental in generalizing the labor market impact of immigration from one country to another. Lastly, differences in labor market rigidity across states are likely important drivers of the large heterogeneity of estimates in the literature.

References

????

Addison, Thomas and Christopher Worswick. 2002. "The Impact of Immigration on the Earnings of Natives: Evidence from Australian Micro Data." *Economic Record* 78 (240):68–78.

Altonji, Joseph G. and David Card. 1991. "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives. Chapter 7 in M. J. Abowd and R. B. Freeman (Eds.)." *Immigration, Trade and the Labor Market* (Chicago: University of Chicago Press):201–234.

Angrist, Joshua D. and Adriana D. Kugler. 2003. "Protective or Counter-Productive? Labour Market Institutions and the Effect of Immigration on EU Natives." *The Economic Journal* 113 (488):302–331.

Avouyi-Dovi, Sanvi, Denis Fougère, and Erwan Gautier. 2013. "Wage rigidity, collective bargaining, and the minimum wage: evidence from French agreement data." *Review of Economics and Statistics* 95 (4):1337–1351.

Barrett, Alan, Adele Bergin, and Elish Kelly. 2011. "Estimating the Impact of Immigration on Wages in Ireland." *The Economic and Social Review* 42 (1):1.

Basten, Christoph and Michael Seigenthaler. 2019. "Do Immigrants Take or Create Residents' Jobs? Evidence from Free Movement of Workers in Switzerland." *Scandinavian Journal of Economics* 121 (3):994–1019.

Bergh, Andreas. 2017. "Explaining the labor market gaps between immigrants and natives in the OECD." *Migration Letters* 14 (2):251–262.

Boeri, Tito and Jan Van Ours. 2013. *The Economics of Imperfect Labor Markets*. Princeton University Press.

Borjas, George, Richard Freeman, and Lawrence Katz. 1996. "Searching for the Effect of Immigration on the Labor Market." *American Economic Review* 86 (2):246–51.

Borjas, George J. 2003. "The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market." *Quarterly Journal of Economics* 118 (4):1359–1374.

———. 2006. "Native Internal Migration and the Labor Market Impact of Immigration." *Journal of Human resources* 41 (2):221–258.

Borjas, George J. 2013. "The Analytics of the Wage Effect of Immigration." *IZA Journal of Migration* 2 (1):22.

Brücker, Herbert, Andreas Hauptmann, Elke J Jahn, and Richard Upward. 2014. "Migration and imperfect labor markets: Theory and cross-country evidence from Denmark, Germany and the UK." *European Economic Review* 66:205–225.

Card, David. 1990. "The Impact of the Mariel Boatlift on the Miami Labor Market." *Industrial and Labor Relations Review* 43 (2):245–257.

———. 2009. "Immigration and Inequality." *American Economic Review: Papers & Proceedings* 99 (2):1–21.

Card, David, Jochen Kluve, and Andrea Weber. 2018. "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations." *Journal of the European Economic Association* 16 (3):894–931.

Card, David and Alan B Krueger. 1995. "Time-Series Minimum-Wage Studies: A Meta-analysis." *The American Economic Review* 85 (2):238–243.

Card, David and Thomas Lemieux. 2001. "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis." *The Quarterly Journal of Economics* 116 (2):705–746.

Card, David and Ethan Lewis. 2007. *The diffusion of Mexican immigrants during the 1990s: explanations and impacts*. Chicago. University of Chicago Press, mexican immigration to the united states ed.

Christensen, Garret and Edward Miguel. 2018. "Transparency, reproducibility, and the credibility of economics research." *Journal of Economic Literature* 56 (3):920–80.

Dustmann, Christian and Albrecht Glitz. 2015. "How Do Industries and Firms Respond to Changes in Local Labor Supply?" *Journal of Labor Economics* 33 (3):711–750.

Dustmann, Christian, Uta Schönberg, and Jan Stuhler. 2016. "The Impact of Immigration: Why Do Studies Reach Such Different Results?" *Journal of Economic Perspectives* 30 (4):31–56.

Edo, Anthony and Hillel Rapoport. 2019. "Minimum Wages and the Labor Market Effects of Immigration." *Labour Economics* 61.

European Commission. 2018. "Eurobarometer." `http://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/Survey/getSurveyDetail/instruments/STANDARD/surveyKy/2180`.

Farber, Henry S, Daniel Herbst, Ilyana Kuziemko, and Suresh Naidu. 2018. "Unions and Inequality Over the Twentieth Century: New Evidence from Survey Data." Tech. rep., National Bureau of Economic Research.

Fleischmann, Fenella and Jaap Dronkers. 2010. "Unemployment among immigrants in European labour markets: an analysis of origin and destination effects." *Work, Employment and Society* 24 (2):337–354.

Foged, Mette and Giovanni Peri. 2015. "Immigrants' Effect on Native Workers: New Analysis on Longitudinal Data." *American Economic Journal: Applied Economics* 8 (2):1–34.

Friedberg, Rachel M. 2001. "The Impact of Mass Migration on the Israeli Labor Market." *Quarterly Journal of Economics* 116 (4):1373–1408.

Glitz, Albrecht. 2012. "The Labor Market Impact of Immigration: A Quasi-Experiment Exploiting Immigrant Location Rules in Germany." *Journal of Labor Economics* 30 (1):175–213.

Hunt, Jennifer. 2017. "The Impact of Immigration on the Educational Attainment of Natives." *Journal of Human Resources* 52 (4):1060–1118.

IDEAS/RePEc. 2020. "Aggregate Rankings for Journals." URL `https://ideas.repec.org/top/top.journals.all.html`.

Kogan, Irena. 2006. "Labor Markets and Economic Incorporation among Recent Immigrants in Europe." *Social Forces* 85 (2):697–721.

Lafortune, Jeanna, Ethan Lewis, and Jose Tassada. 2019. "People and Machines: A Look at The Evolving Relationship between Capital and Skill in Manufacturing, 1860-1930, Using Immigration Shocks." *Review of Economics and Statistics* 101 (1):30–43.

Lewis, Ethan. 2011. "Immigration, Skill Mix, and Capital-Skill Complementarity." *Quarterly Journal of Economics* 126 (2):1029–1069.

———. 2013. "Immigration and Production Technology." *Annual Review of Economics* 5 (1):165–191.

Lewis, Ethan and Giovanni Peri. 2015. "Immigration and the Economy of Cities and Regions." *Handbook of Regional and Urban Economics* 5A.

Longhi, Simonetta, Peter Nijkamp, and Jacques Poot. 2005. "A Meta-Analytic Assessment of the Effect of Immigration on Wages." *Journal of Economic Surveys, Wiley Blackwell* 19 (3):451–477.

———. 2008a. "The Impact of Immigration on the Employment of Natives in Regional Labour Markets: A Meta-analysis." In *Migration and Human Capital*. Edward Elgar, 173–193.

———. 2008b. "Meta-Analysis of Empirical Evidence on the Labour Market Impact of Immigration." *Région et Développement* .

———. 2010a. "Joint Impacts of Immigration on Wages and Employment: Review and Meta-Analysis." *Journal of Geographical Systems* 12 (4):355–387.

———. 2010b. "Meta-Analyses of Labour-Market Impacts of Immigration: Key Conclusions and Policy Implications." *Environment and Planning C: Government and Policy* 28 (5):819–833.

Manacorda, Marco, Alan Manning, and Jonathan Wadsworth. 2012. "The Impact of Immigration on the Structure of Wages: Theory and Evidence from Britain." *Journal of the European Economic Association* 10 (1):120–151.

Moore, Henry Ludwell. 1911. *Laws of Wages: An Essay in Statistical Economics*. Macmillan.

National Academies of Sciences, Engineering, Medicine, and others (NAS). 2017. *The Economic and Fiscal Consequences of Immigration*. National Academies Press.

Nickell, Stephen and Richard Layard. 1999. "Labor market institutions and economic performance." *Handbook of labor economics* 3:3029–3084.

Nunziata, Luca. 2005. "Institutions and wage determination: A multi-country approach." *Oxford Bulletin of Economics and Statistics* 67 (4):435–466.

OECD. 2018. "OECD.Stat Labor Market Policies and Institutions Database." URL `https://stats.oecd.org/`.

Ottaviano, Gianmarco I. P. and Giovanni Peri. 2006. "The Economic Value of Cultural Diversity: Evidence from US Cities." *Journal of Economic Geography* 6 (1):9–44.

———. 2012. "Rethinking the Effect of Immigration on Wages." *Journal of the European Economic Association* 10 (1):152–197.

Peri, Giovanni. 2012. "The Effect of Immigration on Productivity: Evidence from U.S. States." *Review of Economics and Statistics* 94 (1):348–358.

———. 2014. "Do immigrant workers depress the wages of native workers?" *IZA World of Labor* :1–10.

Peri, Giovanni and Chad Sparber. 2009. "Task Specialization, Immigration and Wages." *American Economic Journal: Applied Economics* 1 (3):135–169.

Peri, Giovanni and Vasil Yasenov. 2019. "The labor market effects of a refugee wave synthetic control method meets the mariel boatlift." *Journal of Human Resources* 54 (2):267–309.

Reuters. 2018. "Immigration top issue for U.S. voters, economy a close second: Reuters/Ipsos poll." `https://www.reuters.com/article/us-usa-election-immigration/immigration-top-issue-for-us-voters-economy-a-close-second-reuters-ipsos-poll-idUSKBN1JV31K`.

Rodrik, Dani. 1997. *Has Globalization Gone Too Far?* Peterson Institute for International Economics.

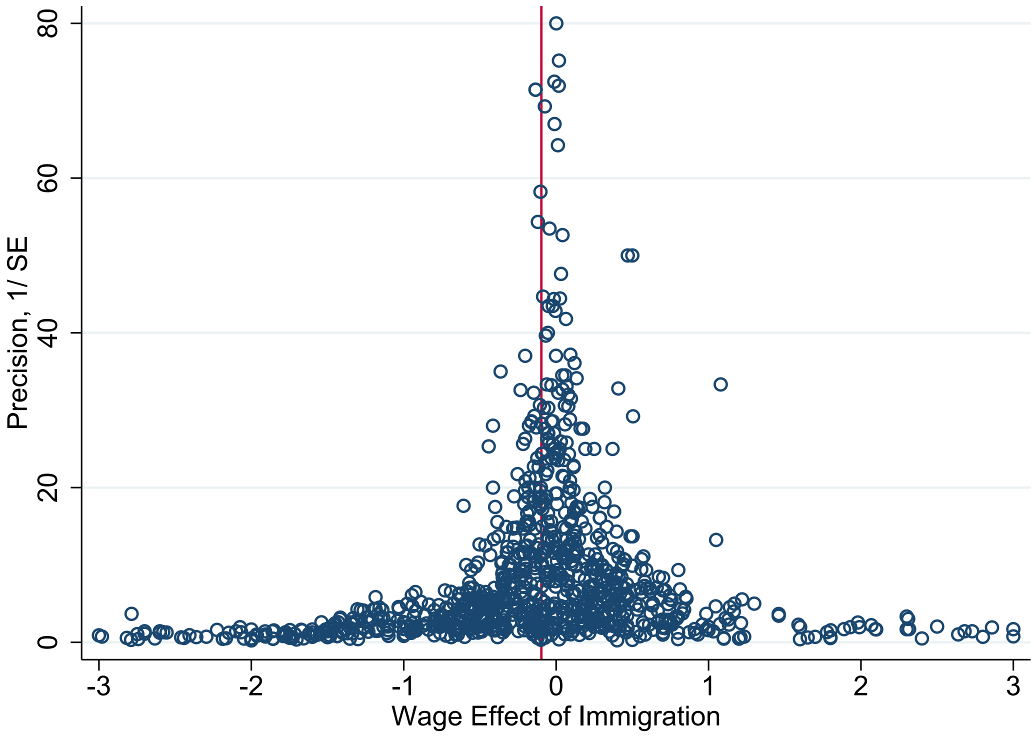
Stanley, Tom D and Hristos Doucouliagos. 2012. *Meta-regression analysis in economics and business*, vol. 5. Routledge.

———. 2014. "Meta-Regression Approximations to Reduce Publication Selection Bias." *Research Synthesis Methods* 5 (1):60–78.

Webber, Douglas. 2014. "Is the Return to Education the Same for Everybody?" *IZA World of Labor* .

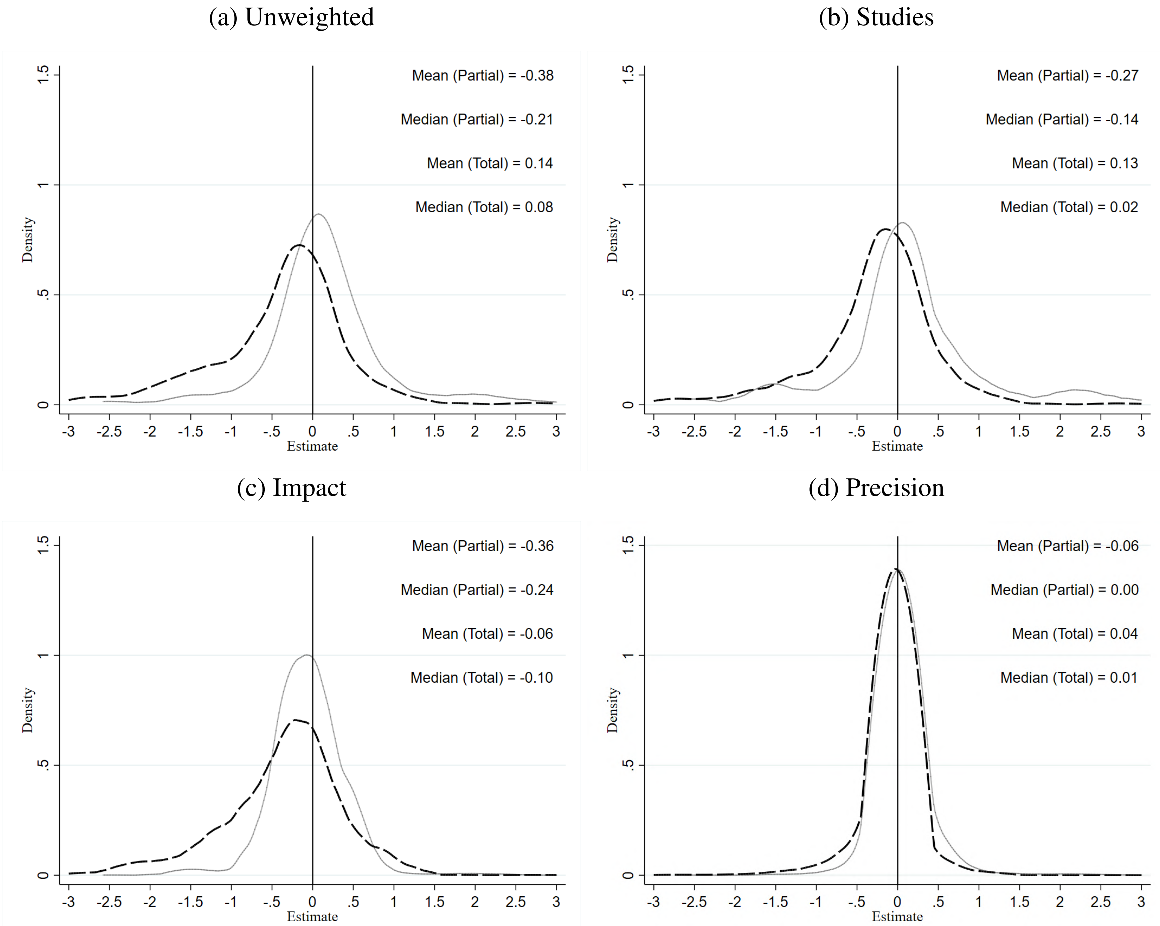
7 Figures and Tables

Figure 1: Publication Bias Test: Funnel Plot



Notes: Scatter plot of precision (inverse standard error) and magnitude of the estimated wage impact magnitudes. The red vertical line shows the median value. A few outliers with precision greater than 80 have been removed for clarity.

Figure 2: Distributions of the Estimates of the Wage Impacts of Immigration



Notes: Kernel densities of partial (dashed) and total (solid) wage effect sizes respectively using the Epanechnikov kernel and Stata's default bandwidth choice. The estimates are unweighted in Panel a, while the rest show weighted distributions: by the inverse number of effect sizes extracted from each study (Panel b), by the inverse journal rank score (Panel c) or by the inverse standard error (Panel d).

Table 1: Institutional Strength and Coverage by Country

	(1) *EPL (Regular)*	(2) *EPL (Collective)*	(3) *Average Job Tenure*	(4) *Collective Bargaining*	(5) *Net Replacement Rate*
			Europe		
Albania	2.14	3.13	-	-	-
Austria	2.61	3.25	10.20	0.97	0.85
Denmark	2.16	3.25	8.17	0.83	0.93
France	2.39	3.38	11.03	0.88	0.85
Germany	2.65	3.63	10.49	0.70	0.92
Ireland	1.40	3.13	9.81	0.40	0.79
Italy	2.76	4.10	12.03	0.80	0.85
Netherlands	2.92	3.01	9.83	0.78	0.80
Norway	2.33	2.50	9.13	0.68	0.86
Portugal	4.52	2.50	12.22	0.76	0.91
Spain	2.75	3.73	9.74	0.80	0.88
Switzerland	1.60	3.63	9.11	0.47	0.90
United Kingdom	1.17	2.86	8.25	0.44	0.65
			North America		
Canada	0.92	2.97	-	0.33	0.85
United States	0.26	2.88	7.51	0.21	0.84
			Rest of the World		
Australia	1.32	2.88	-	0.70	0.65
Israel	2.04	1.88	-	0.52	0.91
Turkey	2.36	2.63	-	0.15	0.75
Mean	2.13	3.11	9.94	0.56	0.83
SD	1.03	0.47	1.36	0.27	0.10

Data source: OECD (2018), job tenure in the United States is from the Job Tenure Supplement of the Current Population Survey.

Notes: Employment protection legislation (EPL) indices range from 0 to 6 (columns 1-2), average job tenure is measured in years (column 3), collective bargaining is the ratio of employees covered by collective agreements divided by all wage earners with the right to bargaining (column 4), and the net replacement rate is the net compensation rate in the initial phase of unemployment (column 5). All values are averaged over the 1998-2016 period. The last two rows show the sample means and standard deviations.

Table 2: Correlation between Wage and Employment Institutions

	(1) *EPL (Regular)*	(2) *EPL (Collective)*	(3) *Average Job Tenure*	(4) *Collective Bargaining*	(5) *Net Replacement Rate*
EPL (Regular)	1.00				
EPL (Collective)	0.05	1.00			
Average Job Tenure	0.73	0.16	1.00		
Collective Bargaining	0.68	0.41	0.59	1.00	
Net Replacement Rate	0.39	0.21	0.29	0.38	1.00

Data source: OECD (2018), job tenure in the United States is from the Job Tenure Supplement of the Current Population Survey.

Table 3: Publication Bias Test: Regression Results

	(1) Unweighted	(2) Studies	(3) Impact	(4) Precision	(5) Combined
Publication Bias Term	9.586	24.941*	-1.479	-36.906	-3.343
	(8.518)	(13.626)	(1.000)	(37.995)	(3.840)
N	981	981	981	981	981
R^2	0.001	0.003	0.000	0.002	0.001

Notes: Each column presents the estimated intercept from a bivariate regression of the t-statistic on precision (inverse standard error) for different weighting schemes on a sample of estimated wage impacts of immigration. The first column is an unweighted regression while the next four ones are weighted: by the inverse number of estimates extracted from the study (Studies), by the inverse journal score (Impact), by the inverse of the standard error of the estimates (Precision), and finally combining all three weights (Combined). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Summary Statistics of Estimates of the Wage Impacts of Immigration

	Panel A: Partial Wage Effects				Panel B: Total Wage Effects			
	Mean	Median	N	Studies	Mean	Median	N	Studies
Total	-0.38	-0.21	613	28	0.14	0.08	417	27
				By Region				
Europe	-0.26	-0.16	265	14	0.13	0.11	299	14
North America	-0.58	-0.41	286	11	0.16	-0.01	117	12
Rest of the World	-0.03	-0.01	62	3	0.11	0.11	1	1
				By Empirical Strategy				
IV (Bartik Type)	-0.17	-0.15	43	7	-0.17	0.20	16	3
IV (Natural Experiment)	-0.21	0.06	11	2	-0.01	-0.11	37	5
IV (Other)	-1.03	-1.30	115	6	0.33	0.24	108	11
Natural Experiment	-0.20	-0.04	58	5	-0.35	-0.02	29	8
OLS	-0.25	-0.16	386	24	0.16	0.10	227	14
				By Native Education Group				
All	-0.40	-0.24	502	26	0.18	0.19	209	19
High Skill	0.16	-0.10	55	10	0.14	0.06	99	9
Low Skill	-0.81	-0.58	56	14	0.06	-0.01	109	12
				By Journal Rank				
Outside Top 50	-0.19	-0.10	284	15	0.18	0.07	263	14
Top 50	-0.55	-0.41	329	13	0.08	0.12	154	13

Notes: Panels A and B show statistics for partial and total wage effects respectively. N denotes the number of estimates, and Studies is the number of studies. The total number of studies by subgroups exceeds the total number of studies in our database (54 articles study wage impacts) since some studies report estimates for multiple subgroups and methods.

Table 5: Labor Market Rigidity and the Wage Effect of Immigration

	Panel A: Partial Wage Effects					Panel B: Total Wage Effects				
	Unweighted	Studies	Impact	Precision	Combined	Unweighted	Studies	Impact	Precision	Combined
EPL (Regular)	0.029	0.330	-0.004	-0.034	0.169	-0.252	-0.538**	-0.198**	-0.121**	-0.178**
	(0.261)	(0.334)	(0.304)	(0.092)	(0.271)	(0.167)	(0.217)	(0.094)	(0.056)	(0.066)
N	443	443	443	429	429	411	411	411	394	394
EPL (Collective)	0.393***	0.501**	0.084	0.362***	0.266**	-0.118	-0.136	-0.313	-0.160**	-0.354**
	(0.091)	(0.179)	(0.167)	(0.079)	(0.114)	(0.303)	(0.875)	(0.187)	(0.066)	(0.150)
N	443	443	443	429	429	411	411	411	394	394
Average Job Tenure	0.297***	0.317**	0.152*	0.218***	0.064	-0.245**	-0.408**	-0.175**	-0.092**	-0.115**
	(0.069)	(0.116)	(0.081)	(0.051)	(0.107)	(0.115)	(0.148)	(0.084)	(0.038)	(0.053)
N	363	363	363	349	349	410	410	410	393	393
Collective Bargaining	0.396	1.156	1.448	0.208	1.294	0.131	1.089	-0.277	-0.107	-0.124
	(0.570)	(0.845)	(1.417)	(0.469)	(0.873)	(0.599)	(1.301)	(0.509)	(0.245)	(0.429)
N	443	443	443	429	429	411	411	411	394	394
Net Replacement Rate	0.054	1.643	-2.500	-0.186	-2.260	-0.547	-0.762	-0.536	-0.526*	-0.752**
	(1.334)	(1.962)	(1.808)	(0.775)	(2.129)	(0.792)	(1.456)	(0.492)	(0.291)	(0.319)
N	443	443	443	429	429	411	411	411	394	394
Region FE	X	X	X	X	X	X	X	X	X	X
Country-Level Controls	X	X	X	X	X	X	X	X	X	X
Study Characteristics	X	X	X	X	X	X	X	X	X	X

Notes: Each entry is an estimated coefficient from a regression of partial (Panel A) or total (Panel B) wage effects of immigration on institution strength and controls. The institutional variables are described in section 3. To alleviate the influence of large outliers, we exclude all estimates larger than 3 in absolute value. The first column is an unweighted regression while the next four columns in each panel are weighted: by the inverse number of estimates extracted from the study (Studies), by the inverse journal score (Impact), by the inverse of the standard error of the estimates (Precision), and finally combining all three weights (Combined). Region FE are dummies for Europe and the rest of the world (North America is the reference). Country-level controls are GDP growth and unemployment rate. Study characteristics are dummies for high- and low-skilled natives (the entire workforce is the reference), dummy for IV (OLS is the reference), and a dummy for annual or more frequent data (less frequent is the reference). Standard errors shown in parentheses are clustered by study. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A Appendix A: Theoretical Background

A.1 Nested CES and Empirical Specifications of Partial and Total Effects

A popular characterization of the workforce distinguishes workers by education and experience and combines them in a nested constant elasticity of substitution (CES) structure (e.g., Card and Lemieux, 2001; Borjas, 2003; Manacorda, Manning, and Wadsworth, 2012; Ottaviano and Peri, 2012).[21] Furthermore, labor and capital are usually combined in a Cobb-Douglas (or a CES) production function. Hence, equations (1) and (2) in section 2 of the main text become

$$Y = AK^{\alpha}L^{1-\alpha}$$

$$L = \left[\sum_{g=1,2,..n} A_g L_g^{\frac{\beta-1}{\beta}} \right]^{\frac{\beta}{\beta-1}}$$

$$L_g = \left[\sum_{a=1,2,..m} A_{ga} L_{ga}^{\frac{\gamma-1}{\gamma}} \right]^{\frac{\gamma}{\gamma-1}}$$

where g and a denotes education levels and age groups, respectively. The parameter β is the elasticity of substitution across education groups and γ is the elasticity of substitution across experience groups. The education (experience) groups are perfect substitutes when β (γ) goes to infinity. Skill cells denoted by s in the main text are denoted by ga combinations, reflecting the two skill dimensions.

The CES structure provides a simple expression of the log marginal productivity of each skill cell as a function of its supply, the aggregate skill supplies and the elasticities of substitution across skill categories.[22] To see this, state the first order condition of the profit maximization problem

$$\begin{aligned} w_{ga} = \frac{\partial Y}{\partial L_{ga}} &= \frac{\partial Y}{\partial L} \frac{\partial L}{\partial L_g} \frac{\partial L_g}{\partial L_{ga}} \\ &= (1-\alpha)AL^{-\alpha}K^{\alpha}A_g L_g^{\frac{-1}{\beta}} L^{\frac{1}{\beta}} A_{ga} L_{ga}^{\frac{-1}{\gamma}} L_g^{\frac{1}{\gamma}} \end{aligned}$$

and take logs and total derivatives to get the unconditional labor demand expressed in percentage changes

$$\frac{dw_{ga}}{w_{ga}} = \alpha \left(\frac{dK}{K} - \frac{dL}{L} \right) + \frac{1}{\beta} \left(\frac{dL}{L} - \frac{dL_g}{L_g} \right) + \frac{1}{\gamma} \left(\frac{dL_g}{L_g} - \frac{dL_{ga}}{L_{ga}} \right)$$

[21] See also Dustmann, Schönberg, and Stuhler (2016) for careful derivation of the national skill approach, the spatial approach (within skill cells) and the combination they label the mixture approach. They use two education and two experience groups to define the three approaches in simple differences. We follow the framework laid out in section 2 and define the approaches as fixed effects models.

[22] See Lewis and Peri (2015) and Dustmann, Schönberg, and Stuhler (2016) for similar derivations.

Hencem, this nested CES model suggests an estimate of the partial own-wage elasticity $\frac{dw_{ga}/w_{ga}}{dL_{ga}/L_{ga}} = -\frac{1}{\gamma} < 0$ can be obtained by simply regressing the change in log wages in skill-cell ga on the change in the cell-specific log labor force while holding the aggregate and the education-specific labor supplies constant by absorbing their changes with fixed effects ($d \log w_{gat} = d\pi_t + (d\pi_t \times s_g) - \frac{1}{\gamma} d \log L_{gat}$). The careful skill-cell specifications are slightly more general, controlling for shocks that are common to all workers such as TFP growth (π_t), shocks that are common to everyone in the same education group ($s_g \times \pi_t$), shocks that are shared within experience groups ($x_a \times \pi_t$) and the education-experience specific productivity ($s_g \times x_a$) in a fixed effect model of the following form:

$$\log w_{gat} = \theta^{skill} p_{gat} + s_g + x_a + \pi_t + (s_g \times x_a) + (s_g \times \pi_t) + (x_a \times \pi_t) + \varphi_{gat} \tag{4}$$

The natural logarithm of L_{gat} is replaced by $p_{gat} = \frac{L^{Im}_{gat}}{L_{gat}}$ in the empirical models and θ^{skill} in equation (4) becomes a semi-elasticity capturing the percentage change in the wage for a one percentage-point change in the immigrant share. This specification follows Borjas (2003) and is by far the most common in the estimation of partial effects representing 12 out of 28 studies and 195 out of 613 estimates classified as providing relative partial effects. An additional 50 estimates from 4 papers provides a similar estimate using occupation-experience cells (e.g, Basten and Seigenthaler, 2019). In total, 18 and 12 papers provide a partial semi-elasticity by using, respectively, education and occupation, sometimes dropping the differentiation by age (e.g, Friedberg, 2001) and/or distinguishing regional labor markets r in the variation in the immigrant share and changing the fixed effects accordingly.[23] Estimates based on variation in some broadly defined skill cells and regions in a "mixture approach" are important contributions to the pool of estimated partial effects. In fact, the second and third most common variation after the classical education-experience-time are occupation-region-time (80 estimates, 5 papers, e.g. Addison and Worswick, 2002) and education-region-time (65 estimates, 5 papers, e.g. Borjas, Freeman, and Katz, 1996).

The total effect on skill cell ga can be obtained by relating wages of workers in this cell to the immigrant share across regions and time, controlling for level-differences across regions (z_{gar}) as well

[23] A few studies further differentiates workers by gender or industry see Appendix Table B.1. These are not published in top journals and hence contribute little when we weight by impact (inverse journal score). Note, Barrett, Bergin, and Kelly (2011) and Glitz (2012) provide estimates based on education as well as occupation and are therefore counted twice here. We have 28 papers providing some type of partial effect.

as aggregate time shocks common to all regions (π_{gat}):[24]

$$y_{gart} = \mu_{ga}^{spatial} p_{rt} + z_{gar} + \pi_{gat} + \varphi_{gart}$$

Many studies relying on spatial variation only report estimates for all workers (or by high and low skilled). We call these average total effects and the estimations take the following general form:

$$y_{gart} = \mu^{spatial} p_{rt} + s_g + x_a + z_r + \pi_t + \varphi_{gart} \tag{5}$$

The important terms are z_r and π_t because this is the level of variation in the immigrant share defining the spatial approach. These regression usually include some control for education and age (s_g and x_a).

We flag 417 estimates from 27 studies as estimating a total wage effect based on spatial variation. A few, mainly old studies, report estimates based on cross-sectional variation only (e.g. Altonji and Card, 1991) and they are included as well. More recent examples of the spatial approach include Foged and Peri (2015); Peri and Yasenov (2019).

A.2 Average Total Effects

For simplicity, consider the case of two broad skill categories, L_1 and L_2. Using the same notation as in section 2, average wages for natives can be expressed as:

$$\begin{aligned}\bar{w}^N &= \frac{1}{L_1^N + L_2^N}\left(L_1^N w_1 + L_2^N w_2\right) \\ &= \frac{G_L}{L_1^N + L_2^N}\left(L_1^N F_1 + L_2^N F_2\right),\end{aligned}$$

where F_1 and F_2 denote the respective partial derivatives. Assume that $\frac{K}{L}$ is a constant.[25] An immigration-induced supply change in one skill group ($dL_1^{Im} > 0$) produces the following impact on the average wages of native workers:

$$\begin{aligned}\frac{\partial \bar{w}^N}{\partial L_1} &= \frac{G_L}{L_L^N + L_2^N}\left(L_1^N F_{11} + L_2^N F_{21}\right) \\ &= \frac{G_L F_{21}}{L_1^N + L_2^N}\left(L_1^N \frac{F_{11}}{F_{21}} + L_2^N\right) \end{aligned} \tag{6}$$

[24]Note, equation (5) is estimated separately for each skill cell ga and hence one obtain an estimate of the total effect for each of skill cell (see Dustmann, Schönberg, and Stuhler, 2016, for a detailed discussion of of this parameter)

[25]This can be justified if immigration is a gradual process, allowing firms to invest in capital and expand production as immigrants come in and create an upward pressure on the return to capital. If the immigration episode is abrupt and unexpected, one may think of the derivations as describing the economy in the long run.

The expression in equation (6) is positive if:[26]

$$
\begin{aligned}
L_1^N \frac{F_{11}}{F_{21}} + L_2^N > 0 \\
L_2^N > L_1^N \frac{-F_{11}}{F_{21}} \\
L_2^N > L_1^N \frac{-F_{11} L\sigma}{F_1 F_2}.
\end{aligned} \tag{7}
$$

Therefore, average wages of native workers could increase in response to e.g. low-skilled immigration if (i) the economy has relatively few low skilled natives (L_1^N), (ii) the complementarity with high-skilled is strong (small σ) and (iii) the decline in the marginal productivity of low-skilled is modest (small $|-F_{LL}|$). Hence, immigration does not need to increases innovation and TFP growth in order to produce positive average effects in the long run. Instead, this may arise from simple complementarities with the native-born workforce.

[26]Notice, that $F_{11} < 0$ and $F_{21} > 0$ and the elasticity of substitution takes the following form $\sigma = \frac{F_1 F_2}{F F_{12}}$ because F is homogeneous of degree one. We also know that F_1 is homogeneous of degree zero when F is homogenous of degree one and hence $F_{11}L_1 + F_{21}L_1 = 0$, also know as the Euler identity.

B Appendix B: Additional Material

B.1 List of Studies

Table B.1: List of Studies Included in Analysis

Authors	Year	Journal	Countries	Data Period	N Wage	N Employment	Variation
Addison and Worswick	2002	The Economic Record	Australia	1982-1996	26	-	ort
Altonji and Card	1991	Immigration, trade, and the labor market	U.S.	1970-1980	20	8	r; rt
Angrist and Kugler	2003	Economic Journal	European Economic Area	1983-1999	-	63	rt
Aydemir and Borjas	2011	Journal of Labor Economics	Canada, U.S.	1960-2001	28	-	gat; gart
Aydemir and Kirdar	2017	European Economic Review	Turkey	1985-1990	-	28	rt
Barrett et al.	2011	The Economic and Social Review	Ireland	1999-2007	30	-	oat; gat
Basten and Siegenthaler	2019	Scandinavian Journal of Economics	Switzerland	2002-2011	4	2	oat
Bauer et al.	2013	Review of International Economics	Germany	2000-2005	19	13	rt
Borjas	2014	Immigration Economics	U.S.	1960-2010	3	-	gat
Borjas	2017	Industrial and Labor Relations Review	U.S.	1977-1992	6	-	rt
Borjas	2003	The Quarterly Journal of Economics	U.S.	1960-2000	28	-	gat; gart
Borjas et al.	1996	American Economic Review	U.S.	1980-1990	20	-	r; grt; gr
Borjas et al.	1997	Brookings Papers on Economic Activity	U.S.	1960-1990	28	14	grt
Bratsberg and Raaum	2012	Economic Journal	Norway	1998-2005	33	-	ot
Bratsberg et al.	2014	Scandinavian Journal of Economics	Norway	1993-2006	31	-	gat
Brunelloa et al.	2020	Journal of Labor Economics	Italy	2006-2016	4	-	rt
Card	1990	Industrial and Labor Relations Review	U.S.	1982-1979	3	3	rt
Card	2001	Journal of Labor Economics	U.S.	1990	14	14	or
Card and Peri	2016	Journal of Economic Literature	U.S.	1960-2010	4	-	gat
Carrasco et al.	2008	Journal of Population Economics	Spain	1991-2001	10	13	gakr; gak; gakt; gakrt
Carrington and De Lima	1996	Industrial and Labor Relations Review	Portugal	1975-1973	7	-	rt
Cattaneo et al.	2015	Journal of Human Resources	Western Europe	1994-2001	14	14	ort
Clark and Drinkwater	2009	Nordic Journal of Political Economy	U.K.	2000-2007	6	6	gat; oat
Cohen-Goldner and Paserman	2011	European Economic Review	Israel	1989-1999	20	20	gat; ot; ort; iot
D'Amuri and Peri	2014	Journal of the European Economic Association	Western Europe	1996-2010	-	22	gart
Dustmann and Glitz	2015	Journal of Labor Economics	Germany	1985-1995	11	-	grt
Dustmann et al.	2005	Economic Journal	U.K.	1983-2000	5	12	rt
Dustmann et al.	2013	Review of Economic Studies	U.K.	1997-2005	66	-	rt

Dustmann et al.	2017	The Quarterly Journal of Economics	Germany	1990-1993	24	-	rt
Edo	2015	The B.E. Journal of Economic Analysis & Policy	France	1990-2002	29	55	gat
Enchautegui	1995	Contemporary Economic Policy	U.S.	1980-1990	8	-	rt
Foged and Peri	2016	American Economic Journal: Applied Economics	Denmark	1995-2008	12	-	rt
Friedberg	2001	The Quarterly Journal of Economics	Israel	1989-1994	16	3	ot; o
Gavosto et al.	1999	LABOUR	Italy	1990-1995	15	-	irt
Glitz	2012	Journal of Labor Economics	Germany	1996-2001	13	10	ort; grt
González and Ortega	2011	Labour Economics	Spain	2001-2006	6	12	grt
Hausmann and Nedelkoska	2018	European Economic Review	Albania	2012-2014	5	-	rt
Hothckiss et al.	2015	Southern Economic Journal	U.S.	1995-2005	10	-	irt
Hunt	1992	Industrial and Labor Relations Review	France	1962-1962	5	4	r; rt
Jean and Jimenez	2011	European Journal of Political Economy	OECD countries	1984-2003	-	24	gart; rt
LaLonde and Topel	1991	Immigration, trade, and the labor market	U.S.	1980	16	-	rt
Lemos and Portes	2014	The B.E. Journal of Economic Analysis & Policy	U.K.	2004-2006	8	36	rt; ot; ort
Llull	2017	Journal of Human Resources	Canada, U.S.	1960-2000	104	-	gart; gat
Mitaritonna et al.	2017	European Economic Review	France	1996-2005	6	-	rt
Monras	2020	Journal of Political Economy	U.S.	1994-1995	4	-	rt
Moreno-Galbis and Tritah	2016	European Economic Review	Western Europe	1998-2004	-	6	ort
Olney	2012	Canadian Journal of Economics	U.S.	2000-2006	20	-	irt
Orrenius and Zavodny	2007	Labour Economics	U.S.	1994-2000	31	-	ort
Ortega and Verdugo	2014	Labour Economics	France	1968-1999	14	18	gart; gat
Pedace	2006	American Journal of Economics and Sociology	U.S.	1980-1990	21	-	r
Pedace	1998	Eastern Economic Journal	U.S.	1980-1990	12	-	r
Peri and Yasenov	2018	Journal of Human Resources	U.S.	1973-1991	5	1	rt
Pischke and Velling	1997	The Review of Economics and Statistics	Germany	1985-1989	-	12	rt
Reed and Danziger	2007	American Economic Review	U.S.	1989-1999	12	12	rt
Schmidt and Jensen	2013	The Annals of Regional Science	Denmark	1997-2006	2	-	rt
Smith	2012	Journal of Labor Economics	U.S.	1980-2007	6	-	rt
Steinhardt	2011	The B.E. Journal of Economic Analysis & Policy	Germany	1975-2001	49	1	oat; gat
Tumen	2016	American Economic Review	Turkey	2010-2013	1	3	rt
Winter-Ebmer and Zweimüller	1999	Journal of Population Economics	Austria	1989-1991	-	3	r
Winter-Ebmer and Zweimüller	1996	Oxford Economic Papers	Austria	1991	4	-	r
Zorlu and Hartog	2005	Journal of Population Economics	Netherlands, U.K., Norway	1989-1998	132	-	r

Notes: Year refers to the year of publication. Data period refers to the overall period of consideration in the study; estimates within a study are sometimes based on sub-periods. N denotes the number of wage and employment effect sizes obtained from each study. The last column shows the variation used to identify the impact of immigration (p in equation 4 and 5) where g stands for education groups, a age groups, t is time, r is region, k gender, i industry and o occupations.

B.2 Robustness Checks

Table B.2: Labor Market Rigidity and the Wage Effect of Immigration, Including |Effect Sizes| > 3

	Panel A: Partial Wage Effects					Panel B: Total Wage Effects				
	Unweighted	Studies	Impact	Precision	Combined	Unweighted	Studies	Impact	Precision	Combined
EPL (Regular)	0.044	0.196	-0.149	-0.013	0.140	-0.299	-0.530**	-0.197**	-0.118**	-0.178**
	(0.294)	(0.334)	(0.466)	(0.105)	(0.267)	(0.204)	(0.232)	(0.093)	(0.056)	(0.066)
N	497	497	497	483	483	425	425	425	408	408
EPL (Collective)	0.313	0.385	-0.431	0.395***	0.268**	-0.239	-0.030	-0.279	-0.157**	-0.357**
	(0.217)	(0.245)	(0.457)	(0.100)	(0.117)	(0.333)	(0.696)	(0.192)	(0.067)	(0.150)
N	497	497	497	483	483	425	425	425	408	408
Average Job Tenure	0.129	0.214	-0.008	0.219***	0.058	-0.199	-0.149	-0.164*	-0.089**	-0.114**
	(0.224)	(0.192)	(0.219)	(0.062)	(0.118)	(0.163)	(0.300)	(0.080)	(0.038)	(0.052)
N	415	415	415	401	401	424	424	424	407	407
Collective Bargaining	-1.556	-0.102	0.707	-0.029	1.176	-0.317	-0.042	-0.267	-0.103	-0.137
	(1.117)	(1.107)	(1.872)	(0.554)	(0.884)	(0.922)	(1.986)	(0.498)	(0.244)	(0.425)
N	497	497	497	483	483	425	425	425	408	408
Net Replacement Rate	1.026	1.896	-3.905	0.135	-2.303	-1.419	-4.161**	-0.521	-0.547*	-0.759**
	(1.746)	(1.481)	(4.203)	(0.838)	(2.209)	(1.208)	(1.825)	(0.507)	(0.270)	(0.318)
N	497	497	497	483	483	425	425	425	408	408
Region FE	X	X	X	X	X	X	X	X	X	X
Country-Level Controls	X	X	X	X	X	X	X	X	X	X
Study Characteristics	X	X	X	X	X	X	X	X	X	X

Notes: Each entry is an estimated coefficient from a regression of partial (Panel A) or total (Panel B) wage effects on institution strength and controls. The institutional variables are described in section 3. The first column is an unweighted regression while the next four columns in each panel are weighted regressions: by the inverse number of estimates extracted from the study (Studies), by the inverse journal score (Impact), by the inverse of the standard error of the estimates (Precision), and finally combining all three weights (Combined). Region FE are dummies for Europe and the rest of the world (North America is the reference). Country-level controls are GDP growth and unemployment rate. Study characteristics are dummies for high- and low-skilled natives (estimates based on the entire workforce is the reference), dummy for IV (OLS is the reference), and a dummy for annual or more frequent data (less frequent is the reference). Standard errors shown in parentheses are clustered by study. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

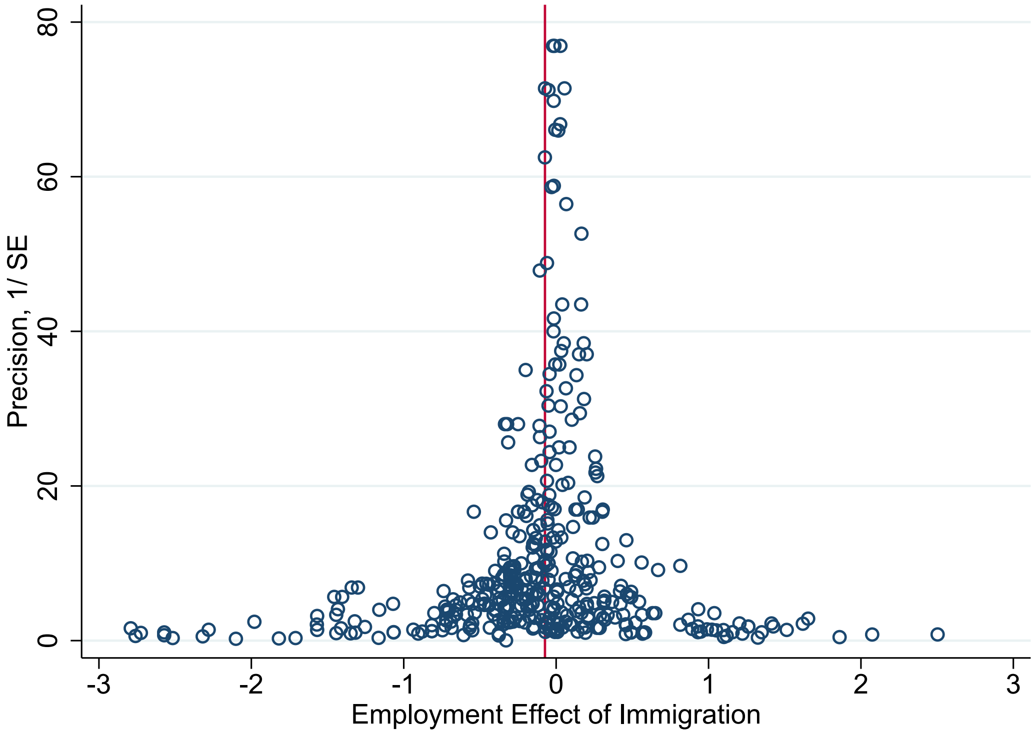
Table B.3: Institutions and the Wage Effect of Immigration on Natives, Robustness Checks

	Panel A: Partial Wage Effects						Panel B: Total Wage Effects					
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
EPL (Regular)	0.169	0.210	-0.389***	0.163	-0.109	0.016	-0.178**	-0.178**	-0.187***	-0.226***	-0.189***	-0.007
	(0.271)	(0.274)	(0.125)	(0.262)	(0.104)	(0.049)	(0.066)	(0.066)	(0.050)	(0.073)	(0.059)	(0.031)
EPL (Collective)	0.266**	0.175	0.431***	0.240**	0.195*	-0.095	-0.354**	-0.354**	-0.405***	-0.478***	-0.247	-0.069
	(0.114)	(0.105)	(0.115)	(0.115)	(0.107)	(0.107)	(0.150)	(0.150)	(0.125)	(0.157)	(0.160)	(0.111)
Average Job Tenure	0.064	-0.145	-0.013	0.125	0.096	0.005	-0.115**	-0.115**	-0.154**	-0.144**	-0.138**	-0.010
	(0.107)	(0.133)	(0.103)	(0.091)	(0.119)	(0.046)	(0.053)	(0.053)	(0.061)	(0.054)	(0.057)	(0.026)
Collective Bargaining	1.294	1.290	-1.158***	1.279	0.106	0.108	-0.124	-0.124	-0.282	-0.278	-0.153	0.085
	(0.873)	(0.857)	(0.273)	(0.832)	(0.538)	(0.254)	(0.429)	(0.429)	(0.412)	(0.471)	(0.702)	(0.175)
Net Replacement Rate	-2.260	-1.399	-0.444	-2.426	-1.863	0.138	-0.752**	-0.752**	-0.897***	-1.079***	-1.038***	-0.607*
	(2.129)	(1.924)	(1.923)	(2.274)	(2.019)	(0.780)	(0.319)	(0.319)	(0.249)	(0.343)	(0.342)	(0.308)
Region FE	X	X	X	X	X	No	X	X	X	X	X	No
Country-Level Controls	X	X	X	X	No	X	X	X	X	X	No	X
Study Characteristics	X	X	X	No	X	X	X	X	X	No	X	X

Notes: Models are weighted linear regressions with the effect size as the dependent variable. Weights correspond to the combined weights used in Table 5. The first column in each panel replicates results from Table 5. Column 2 excludes effect sizes based on variation across industries. Column 3 excludes effect sizes estimated with OLS if IV estimates are reported within studies. Column 4 shows estimates without study characteristics as controls. Column 5 shows estimates from regressions without country-level controls (unemployment rate and GDP growth). Column 6 shows estimates without region FE. Country-level controls and study characteristics are described in Table 5. Standard errors in parentheses are clustered by studies. See the Online Data Appendix for a complete description of all the variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

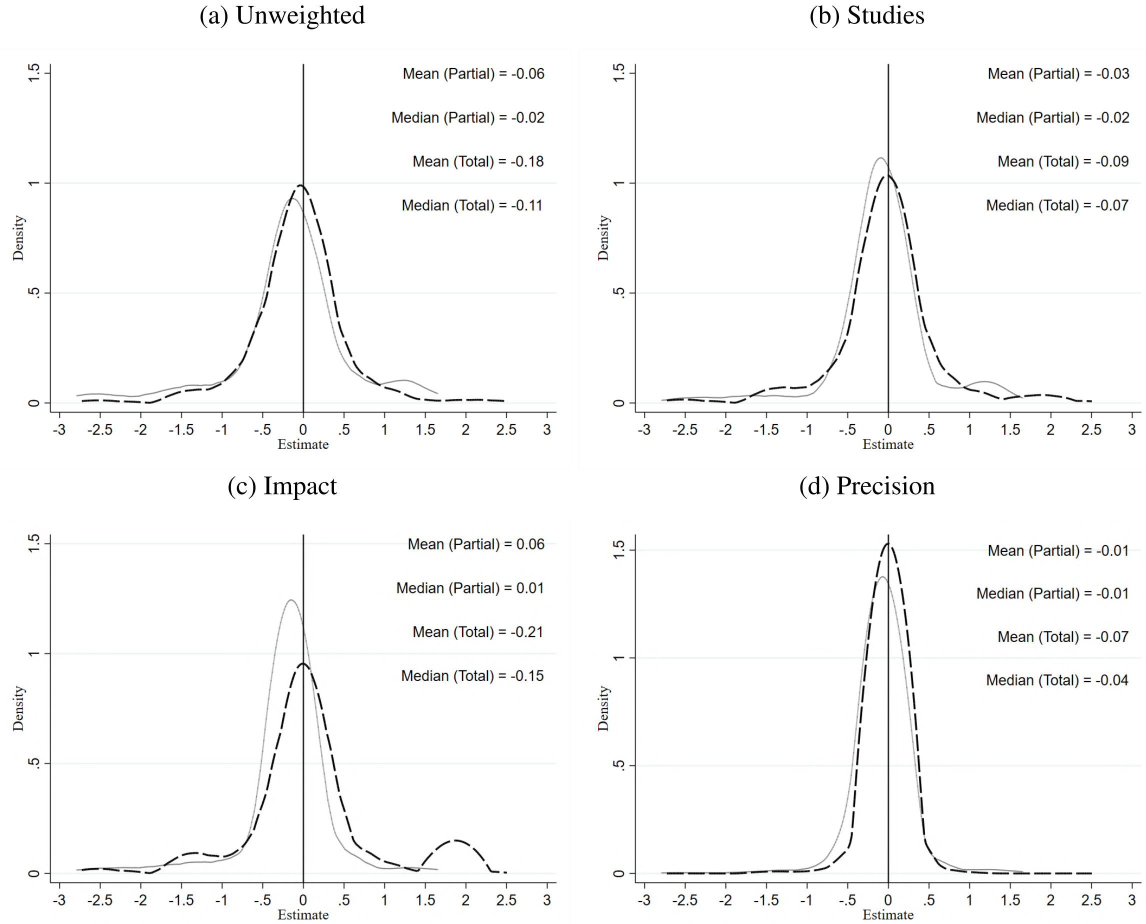
B.3 Analysis of the Employment Effects of Immigration

Figure B.1: Publication Bias Test: Funnel Plot (Employment Effects)



Notes: Scatter plot of precision (inverse standard error) and magnitude of the estimated employment impact magnitudes. The red vertical line shows the median value. A few outliers with precision greater than 80 have been removed for clarity.

Figure B.2: Distributions of Employment Effects of Immigration



Notes: Kernel densities of partial (dashed) and total (solid) employment effect sizes respectively using the Epanechnikov kernel and Stata's default bandwidth choice. The estimates are unweighted in Panel a, while the rest show weighted distributions: by the inverse number of effect sizes extracted from each study (Panel b), by the inverse journal rank score (Panel c) or by the inverse standard error (Panel d).

Table B.4: Publication Bias Test: Regression Results (Employment Effects)

	(1) Unweighted	(2) Studies	(3) Impact	(4) Precision	(5) Combined
Publication Bias Term	-1.649	1.669	-5.350**	-1.052	-4.710**
	(3.777)	(5.496)	(2.610)	(2.233)	(2.028)
N	482	482	482	482	482
R^2	0.003	0.002	0.036	0.003	0.009

Notes: Each column presents the estimated intercept from a bivariate regression of the t-statistic on precision (inverse standard error) for different weighting schemes on a sample of estimated employment impacts of immigration. The first column is an unweighted regression while the next four ones are weighted: by the inverse number of estimates extracted from the study (Studies), by the inverse journal score (Impact), by the inverse of the standard error of the estimates (Precision), and finally combining all three weights (Combined). $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

Table B.5: Summary Statistics of Estimated Employment Effects of Immigration

	Panel A: Partial Employment Effects				Panel B: Total Employment Effects			
	Mean	Median	N	Studies	Mean	Median	N	Studies
Total	-0.06	-0.02	222	17	-0.18	-0.11	210	14
				By Region				
Europe	-0.06	-0.02	165	12	-0.18	-0.07	137	7
North America	-0.23	-0.16	28	2	-0.07	-0.09	24	4
Rest of the World	0.05	-0.01	29	3	-0.22	-0.28	49	3
				By Empirical Strategy				
IV (Bartik Type)	0.22	0.20	45	6	-1.07	-1.27	4	1
IV (Natural Experiment)	-0.75	-1.08	4	2	-0.80	-1.07	28	2
IV (Other)					-0.18	-0.12	40	7
Natural Experiment	-0.31	-0.04	35	4	-0.12	-0.12	10	4
OLS	-0.07	-0.05	138	13	-0.01	-0.08	128	10
				By Native Education Group				
All	-0.07	-0.02	204	17	-0.22	-0.12	155	11
High Skill	-0.13	-0.15	7	3	0.08	0.01	15	3
Low Skill	0.05	0.19	11	5	-0.09	-0.09	40	7
				By Journal Rank				
Outside Top 50	-0.07	-0.07	119	9	0.02	-0.06	79	7
Top 50	-0.06	0.00	103	8	-0.30	-0.15	131	7

Notes: Panels A and B show statistics for partial and total employment effects respectively. N denotes the number of estimates, and Studies is the number of studies. The total number of studies by subgroups exceeds the total number of studies in our database (54 articles study wage impacts) since some studies report estimates for multiple subgroups and methods.

Table B.6: Institutions and the Employment Effect of Immigration

	Panel A: Partial Employment Effects					Panel B: Total Employment Effects				
	Unweighted	Studies	Impact	Precision	Combined	Unweighted	Studies	Impact	Precision	Combined
EPL (Regular)	-0.160	-0.119	-0.714**	-0.143	-0.185	0.237	0.125	-0.002	0.131	-0.067
	(0.242)	(0.254)	(0.281)	(0.188)	(0.234)	(0.144)	(0.130)	(0.204)	(0.174)	(0.162)
N	170	170	170	169	169	128	128	128	125	125
EPL (Collective)	0.048	0.209	-0.954	-0.049	-0.069	0.520	0.267	-0.030	0.305	-0.369
	(0.530)	(0.397)	(0.734)	(0.139)	(0.107)	(0.301)	(0.341)	(0.428)	(0.545)	(0.295)
N	170	170	170	169	169	128	128	128	125	125
Average Job Tenure	-0.117	-0.081	-0.256**	-0.066	-0.215	0.186**	0.092	0.139	0.091	-0.072
	(0.092)	(0.110)	(0.112)	(0.059)	(0.138)	(0.078)	(0.075)	(0.125)	(0.129)	(0.136)
N	147	147	147	146	146	97	97	97	94	94
Collective Bargaining	-0.430	-0.486	0.400	-0.253	-0.170	0.825	0.373	0.165	0.402	0.369
	(0.601)	(0.737)	(1.613)	(0.302)	(0.659)	(0.612)	(0.226)	(0.797)	(0.376)	(0.229)
N	170	170	170	169	169	128	128	128	125	125
Net Replacement Rate	-0.881	0.047	-4.888**	-0.267	-0.316	1.416	0.747	-0.052	0.824	-0.757
	(1.375)	(1.168)	(2.004)	(0.448)	(0.418)	(0.831)	(0.865)	(1.189)	(1.264)	(0.952)
N	170	170	170	169	169	128	128	128	125	125
Region FE	X	X	X	X	X	X	X	X	X	X
Country-Level Controls	X	X	X	X	X	X	X	X	X	X
Study Characteristics	X	X	X	X	X	X	X	X	X	X

Notes: Each entry is an estimated coefficients from a regression of partial (Panel A) or total (Panel B) employment effects of immigration on institution strength and controls. The institutional variables are described in section 3. To alleviate the influence of large outliers, we exclude all estimates larger than 3 in absolute value. The first column is an unweighted regression while the next four columns in each panel are weighted: by the inverse number of estimates extracted from the study (Studies), by the inverse journal score (Impact), by the inverse of the standard error of the estimates (Precision), and finally combining all three weights (Combined). Region FE are dummies for Europe and the rest of the world (North America is the reference). Country-level controls are GDP growth and unemployment rate. Study characteristics are dummies for high- and low-skilled natives (the entire workforce is the reference), dummy for IV (OLS is the reference), and a dummy for annual or more frequent data (less frequent is the reference). Standard errors shown in parentheses are clustered by study. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.