



HAL
open science

Essays on migration and productivity

Cem Ozguzel

► **To cite this version:**

Cem Ozguzel. Essays on migration and productivity. Economics and Finance. Université Panthéon-Sorbonne - Paris I, 2019. English. NNT : 2019PA01E053 . tel-03381203

HAL Id: tel-03381203

<https://theses.hal.science/tel-03381203>

Submitted on 16 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Essays on Migration and Productivity

Cem Özgüzel

*Thèse pour l'obtention du grade de Docteur de l'Université Paris 1 Panthéon-Sorbonne et de
l'École d'Économie de Paris en Sciences Économiques.*

Présentée et obtenue publiquement à Paris le 19 novembre 2019.

Sous la direction de Hillel Rapoport

Professeur à l'Université Paris 1 Panthéon-Sorbonne et PSE

et co-direction de Ariell Reshef

Directeur de recherche CNRS, Paris Panthéon-Sorbonne et PSE

Rapporteur	Jesús Fernández-Huertas Moraga	Professeur, Université Carlos III de Madrid
Rapporteur	Gregory Verdugo	Professeur, Université Evry Val-d'Essonne
Examineur	Pierre-Philippe Combes	Directeur de recherche CNRS, GATE Lyon & Sciences Po
Présidente	Maria Bas	Professeure, Université Paris 1 Panthéon-Sorbonne

Acknowledgements

As with everything in life, this dissertation would not have been possible without the generous and constant support of many people.

I want to express my sincere gratitude to my supervisor Hillel Rapoport for his continuous encouragement and guidance. Hillel's natural curiosity, coupled with boundless knowledge, helped me understand the complexities, impact, and rewarding nature of research. I will always be grateful for these invaluable lessons.

I would like to thank Ariell Reshef, my co-supervisor, for his continued support and mentorship throughout this process. Everyday, Ariell encouraged me to push my capacity frontier further by refusing to settle for less. While insisting on rigor and quality, Ariell made sure to provide generous support and commitment in helping me navigate uncharted waters.

I would also like to extend my sincere gratitude to Pierre-Philippe Combes, for generously sharing his extensive knowledge and supervising the chapter on Turkey. I thank him for his invaluable support, warmth and guidance.

I want to thank referees Jesús Fernández-Huertas Moraga and Gregory Verdugo for their generous guidance and support throughout the final stretch of this dissertation and over my doctoral studies. I would like to thank Prof. Maria Bas for her valuable feedback and presiding over the committee.

This dissertation benefited from the constant guidance of many members of the PSE and Paris 1 faculty. My sincere thanks goes to Angelo Secchi, for his mentorship from masters until the very end of my PhD. I will forever be grateful for his unlimited patience and smile.

I would also extend my gratitude to François Fontaine and David Margolis for sharing their extensive knowledge and curiosity; Simone Bertoli, Frédéric Docquier and Anthony Edo for their helpful feedback throughout the years. I would like to thank my co-authors Dany Bahar and Andreas Hauptmann, for the opportunity to work as a team and learn from them.

I would like to thank Ricardo Hausmann, Frank Neffke and Ljubica Nedelkoska, for their inspiration and for providing the opportunity to spend six months in an academic context few get to experience. I would also like to thank, Herbert Bruecker for inviting me to IAB and letting me be part of his team;

with a special thanks to Ignat Stepanok, Stella Capuano and Ehsan Vallizadeh for their warm welcome. Nurnberg will always have a special place in my heart.

I would like to thank Mouez Fodha and Sylvie Lambert for providing all of us their warm support and constant effort to provide us the best conditions possible through the challenging years of PhD. Special thanks to Véronique Guillotin, Jean-Christophe Fouillé, Loïc Sorel, Beatrice Havet for their professionalism and care.

A special thanks to my colleagues and friends from MSE for making me feel part of a research unit and helping me develop as a researcher: François Facchini, Pierre Kopp, Carine Staropoli, Patricia Vornetti, Sophie Massin, Michael Melki, Milena Spach, Antoine Pietri, Riham Ahmed Ezzat, Eric Dubois, Maxime Fajeau, Alexandre Mayol; and in particular, my colleague, roommate, and good friend Benjamin Michallet.

I would like to thank my colleagues and friends at PSE and Paris 1 who have shared in the daily struggles of a PhD and made every day at the office feel like anything but work: Viviane Azais, Alexia Lochmann, Sulin Sardoschau, Alexandra Jarotschkin, Elias Bouacida, Cristina Herghelegiu, Evgenii Monastyrenko, Marco Ranaldi, Emanuele Franceschi, Guillaume Claveres, Pauline Wibaux, Jaime Leyva, Jaime Montana Doncel, and Nhung Luu.

A special thanks to my friends for their generous help at various stages of this dissertation: Luis Estevez Bauluz, Sebastien Bock, Güney Celbiş, José Montalbán Castilla, Martin Fernández-Sánchez, Jared Gars, Davide Luca, Elisabetta Pasini and Seyhun Orcan Sakalli.

Pursuing a PhD is a unique and at times isolating experience, yet within the ambiguity and fog, one benefits greatly from the warmth and support of fellow colleagues who also become part of the experience. Sharing a challenging experience and providing support to one another creates strong bonds. I would like to thank in particular Niels Boissonnet, Paul Brandily-Snyers, Michelle Marshalian and Hector Moreno for their fraternity and solidarity until the very end.

I would also like to thank my friends outside of PhD who throughout these years extended me their support, and shared their excitement about my work. I would like to thank Sinan Açıkgöz, Roberto Cardinale, Safae Derj, Nicolas Denis, Emir Han Hacıoğlu, Beatrice Garapon, Herve Gonçalves, Selin Güler, Pinar Güven, Nadeen Hussein, Romain Lecler, Ecem Okan, Yiğit Şakar and Arif Yıldız. A very special thanks go to Garrett Brady and Thomas Dournon.

Finally, and most importantly, I would like to thank my family, Can, Mine and Hayri for their unlimited support and unconditional love. I dedicate this dissertation to them.

To Can, Mine and Hayri

Contents

1	General Introduction	8
1.1	Cushioning Effect of Immigrant Mobility: Evidence from the Great Recession	16
1.2	Migration and Post-conflict Reconstruction: The Effect of Returning Refugees on Export Performance in the Former Yugoslavia	18
1.3	Agglomeration Effects in a Developing Economy: Evidence from Turkey	20
1.4	Concluding Remarks	21
2	Introduction Generale	22
2.1	Effet Amortisseur de la Mobilité des Immigrants	32
2.2	Migration et Reconstruction Post-conflit : L'Effet du Retour des Réfugiés sur les Export- tations en ex-Yougoslavie	34
2.3	Economies d'Agglomération dans un Economie en Développement : Le Cas de la Turquie	36
3	The Cushioning Effect of Immigrant Mobility: Evidence from the Great Recession in Spain	38
3.1	Introduction	39
3.2	An Equilibrium Model with Heterogeneous Labor Supply, Demand Shock and Wage Rigidities	42
3.2.1	Set-up	43
3.2.2	Equilibrium Effect of Migration and Demand	45
3.3	Spanish Context	47
3.3.1	The Great Recession	47
3.3.2	Immigrant Population in Spain	49
3.3.3	Immigrant vs. Native Mobility	54
3.4	Data and Summary Statistics	56
3.4.1	Data	56
3.4.2	Summary Statistics	61
3.5	Empirical Strategy	62
3.5.1	Econometric Equation	62
3.5.2	Identification	64
3.6	Main Results	68
3.6.1	First Stage and Validity of Instruments	68
3.6.2	Second Stage	70
3.7	Robustness	71
3.7.1	Previous-trends	72

3.7.2	Bartik and Other Controls	74
3.7.3	Alternative Instruments	76
3.7.4	Alternative Measures of the Supply Shock	79
3.7.5	Alternative Weights	81
3.8	Dynamics of Adjustment	82
3.9	Heterogeneity Analysis	83
3.9.1	By Skill Group	83
3.9.2	By Demographics	85
3.9.3	EU versus non-EU outflows	86
3.10	Mechanism	87
3.10.1	Geographic Mobility of Natives	87
3.10.2	Margins of the Employment Effects	90
3.10.3	Margins of the Wage Effects	91
3.10.4	Margins by Contract Types	93
3.11	Understanding the Magnitude of the Estimates	95
3.12	Concluding Remarks	96
3.13	Appendix	97
3.13.1	Derivation of the Firm's Demand Curve	97
3.13.2	Derivation of Labor Supply	99
3.13.3	Derivation of Equilibrium Wage and Employment Responses under Flexible Wages	99
3.13.4	Firms Adjustment During the Crisis	102
3.13.5	Differences in Unemployment Rates	103
3.13.6	Reconciling with Cadena and Kovak (2016)	104
3.13.7	Immigrant Population	108
3.13.8	Government Measures	114
3.13.9	Net Immigrant Departures as a Share of Total Working-age Immigrant Population	115
3.13.10	More on Mobility Data	116
3.13.11	Construction of the instrument	118
3.13.12	Instrument Validity	121
3.13.13	Construction of Bartik control	124
3.13.14	Immigrant and Native Outflows	125
3.13.15	Maps	126
3.13.16	Differentiating Between Boom and Bust	127
3.13.17	Part-time Employment	129

4 Migration and Post-conflict Reconstruction:

The Effect of Returning Refugees on Export Performance in the Former Yugoslavia 130

4.1	Introduction	131
4.2	Historical Context	134
4.3	Labor Market Conditions and Mobility of Refugees	135
4.4	End of the War and Deportation	136
4.5	Data and Sample	137
4.6	Natural Experiment: Yugoslavian Refugees	138
4.6.1	Empirical Strategy	140
4.6.2	Summary Statistics	149
4.7	Results	150
4.7.1	Pre-trend and Event-study Estimation	153
4.7.2	Robustness and Heterogenous Effects	157
4.8	Mechanisms: Heterogeneous Effects by Workers' Characteristics	169
4.9	Concluding Remarks	175
4.10	Appendix	177
4.10.1	Anecdotal Evidence: Four Individual Stories	177
4.10.2	Details on Employment Data, Sample Construction and Variable Description	179
4.10.3	Zeros in the Data	182
4.10.4	Including Exports to Germany	184
4.10.5	Additional Controls	185
4.10.6	Further Tests on the Exogeneity of Exit From the Labor Force	188
4.10.7	Robustness: German WZ 93 3-digit Industries	189
4.10.8	Other Robustness Checks	191
4.10.9	Excluding Slovenia and Macedonia from the Sample	192
4.10.10	Occupations by Characteristics	193
4.10.11	Estimations Using Treatments by Educational Attainment and Occupations Characteristics	195
4.10.12	Expanding to All Countries: External Validation	202

5 Agglomeration Effects in a Developing Economy:

Evidence from Turkey 205

5.1	Introduction	206
5.2	Literature Review	209
5.2.1	Agglomeration Economies	209
5.2.2	Magnitudes for the Effect of Density on Workers' Productivity	210

5.2.3	Evidence from Developing Countries	211
5.3	The Turkish Context	212
5.4	Literature on Regional Differences	215
5.5	Data and Sample	216
5.5.1	Social Security Data	216
5.5.2	Household Labor Force Survey	216
5.5.3	Historical Data	217
5.5.4	Controls	217
5.5.5	Sample	218
5.6	Empirical Strategy	218
5.6.1	Econometric Equation	219
5.6.2	Estimation Issues	219
5.6.3	Estimation Issue 1: Omitted Variables	220
5.6.4	Estimation Issue 2: Circular Causality	224
5.6.5	Estimation Issue 3: Accounting for Local Shocks	225
5.6.6	Estimation Issue 4: Sorting by Ability	226
5.7	Main Results	227
5.7.1	Density	227
5.7.2	Omitted Variable Bias	227
5.7.3	Reverse Causality	229
5.7.4	Sorting By Ability	233
5.8	Multivariate Approach: A Unified Framework	237
5.9	Infrastructure and Other Amenities	240
5.10	Concluding Remarks	242
5.11	Appendix	244
5.11.1	Summary Statistics	244
5.11.2	Yearly OLS results	245
5.11.3	One-step results	246
5.11.4	Individual Analysis	247
5.11.5	Informal Employment	249
5.11.6	First Stage of Multivariate Regression	251
5.11.7	Additional Results: Infrastructure and Amenities	253

1 General Introduction

Migration and productivity are two intertwined and driving forces of our society, as people move to different places seeking better opportunities and improvement of their living standards. The possibility of achieving these goals depends on an individual's characteristics but also those of the place where the individual lives and works. That is why people migrate to locations that have higher productivity as they are more likely to provide the opportunity for economic growth and better living standards.

Starting from the industrial revolution, as the western countries overcame pre-modern growth constraints, productivity and income levels between countries started diverging (Pomeranz, 2001). The increase in inequalities were later followed by the rise of further inequalities between regions within countries. Across history, these spatial differentials generated both internal and international migration. The rise in international migration over the past decades and the recent influx of refugees across Europe has re-centered international migration at the core of both political and academic debate.

In the past decade, these growing trends have placed greater importance on understanding productivity from governments and academics alike. Specifically, decline in productivity growth since the mid-2000s expanded in the aftermath of the Great Recession. The slowdown in productivity growth generated a lively debate on the future of productivity growth across many academic circles. In parallel, in many OECD countries, spatial differences in productivity and income have further widened, creating a divergence in economic outcomes. These developments have revived interest in understanding the determinants of productivity, with a primary focus place on the sources of spatial differences and the causes of the recent slowdown.

This dissertation explores the interaction between migration and productivity, through multiple an-

gles, across three different country and period contexts. Specifically, I study the labor market benefits of immigrant mobility during an economic crisis and its impact on the adjustment of labor markets; diffusion of productive knowledge through migrant mobility across countries and its benefits in the reconstruction of a country in the aftermath of a war; and productivity gains associated with concentration of people in larger urban areas in a developing economy.

Productivity: Why does it matter?

Productivity is the ability to produce more output by better-combining inputs via new ideas, technological innovations, and business models. Innovations such as the steam engine, electrification, and digitization have led to radical changes in the production of goods and services, thus productivity. Productivity is important as it is a primary determinant of wages, which strongly influence material conditions, living standards, and well-being of people. Indeed, historical data show that wages and productivity growth have moved in lock-step over the past millennium. Today the large differences in income per capita observed across countries are primarily reflected by differences in labor productivity (1). Thus productivity growth will remain a primary driver of both economic growth and improvement of individual well-being.

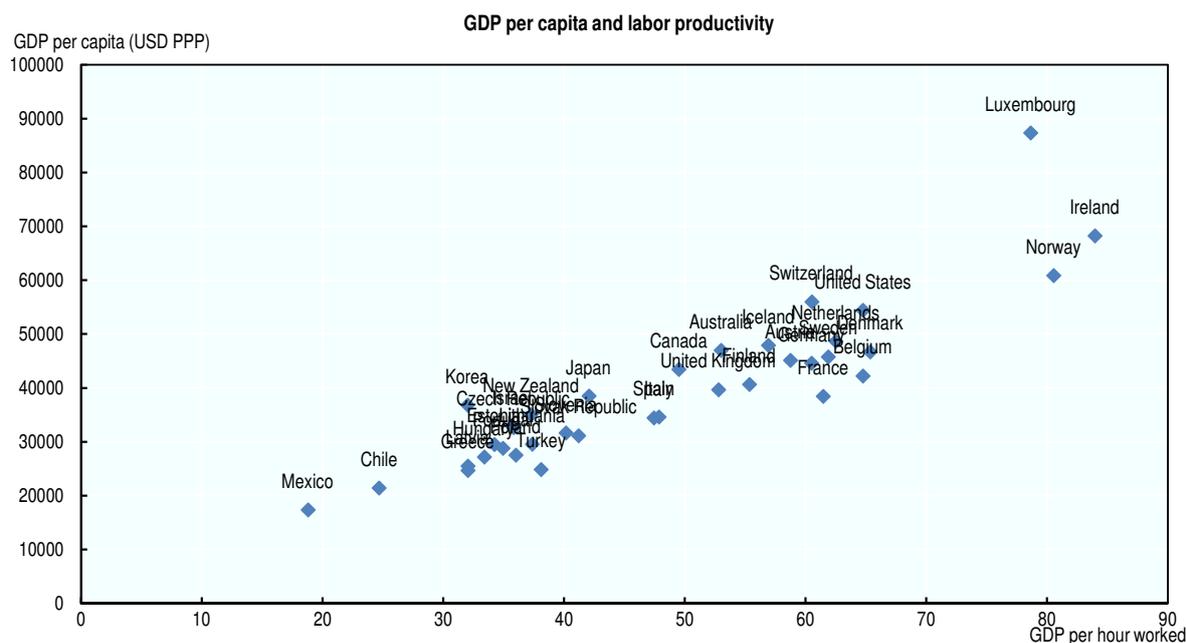
Productivity growth has slowed down in the past two decades

The literature offers various explanations for the productivity slowdown.¹ One of the explanations put forth is the breakdown of the so-called “diffusion machine.” According to this argument top firms (or those at the productivity “frontier”) have continued to innovate and increase their productivity. The productive knowledge generated in these firms, however, has stopped diffusing to other less productive firms generating productivity differences within countries and industries.

Although a clear explanation for this diffusion challenge is yet to be found, there is ample evidence

¹For further evidence on the nature and causes of the productivity slowdown, see Syverson (2017); Cetto et al. (2018)).

Figure 1: Income and productivity



Note: Figure plots GDP per capita in USD (PPP) and GDP per worker in 2018.
Source: OECD Database

that geography plays a substantial role. In many OECD countries, the productivity growth observed since the 2000s was generated by only a few regions, whereas the remaining regions recorded either no growth or even declining productivity. As a result, many OECD countries have experienced substantial increases in the interregional productivity and economic disparities.² While gaps in GDP per capita across OECD countries have narrowed over the last two decades, countries are faced with increasing income gaps within their own regions and cities.

Persistent disparities in economic performance across regions have created a divergence in economic and social outcomes within countries. Poor economic performance, persistently high unemployment, and declining wages diminish their communities' hopes and abilities to achieve a better future. Consequently, in many countries, the growing share of the population is becoming discontented with the status quo. Not surprisingly, the discontent has a geographical pattern and is highly correlated with the economic performance of regions (Rodríguez-Pose, 2018; McCann, 2019).

²Trade shocks and increased competition in manufacturing from developing countries is another explanation put forth for the divergence of productivity and income levels across regions in some of the developed countries. For a review of the literature on the labor market effects of Chinese trade competition see for instance Autor et al. (2016).

Cities: Source of productivity growth

Cities are a key driver of their regions' economic performance and productivity growth. In today's digital and service economy, cities enable interactions between individuals that generate growth through transformative and knowledge-intensive industries.³ The close proximity of economic agents generates knowledge and fosters technological spillover as a consequence of the interactions of workers and their mobility across firms. While innovation can happen anywhere, it occurs mainly in highly urbanized areas due to strong spatial decay of knowledge (Carlino and Kerr, 2015). Productivity growth generated in cities also spillover to the surrounding regions and the rest of the economy impacting their economic performance. Thus by expanding the productivity frontier and generating productivity growth, cities are essential drivers of their region's and country's potential for long-run economic growth. Urbanization in developing countries plays an important role in the structural transformation and economic growth of emerging countries. Cities have become economic hubs, pioneering the rest of the economy in terms of innovation and productivity but also in terms of cultural, educational, and health amenities. The strong economic magnet, coupled with attractive amenities, continues to attract internal migration from rural areas, which is the main force of urban population growth in many developing countries. As a result, although these cities enjoy high productivity levels and sometimes quality of life, most of these cities are then tasked with handling the influx of migrant workers. In many developing countries, the absence of adequate institutional and physical infrastructure to absorb a large influx of populations increases housing costs and causes inefficient land-use which dampen the gains associated with denser and larger cities. Understanding the importance of agglomerations in developing countries is valuable because, beyond their size and weight in the overall economy, they are at the heart of structural change and remain the primary destination for internal migrants who seek better lives.

³Productivity gains due to agglomeration of people and firms are not limited to knowledge spillovers. Economist Alfred Marshall (Marshall 1890) was among the first to emphasize that the agglomeration of people and firms can increase productivity. Since then, considerable effort has been spent on identifying the mechanisms through which agglomeration economies emerge. While not all aspects related to emergence of agglomeration economies are fully understood, the benefits can be broadly split into three groups: sharing, matching and learning (Duranton and Puga, 2004).

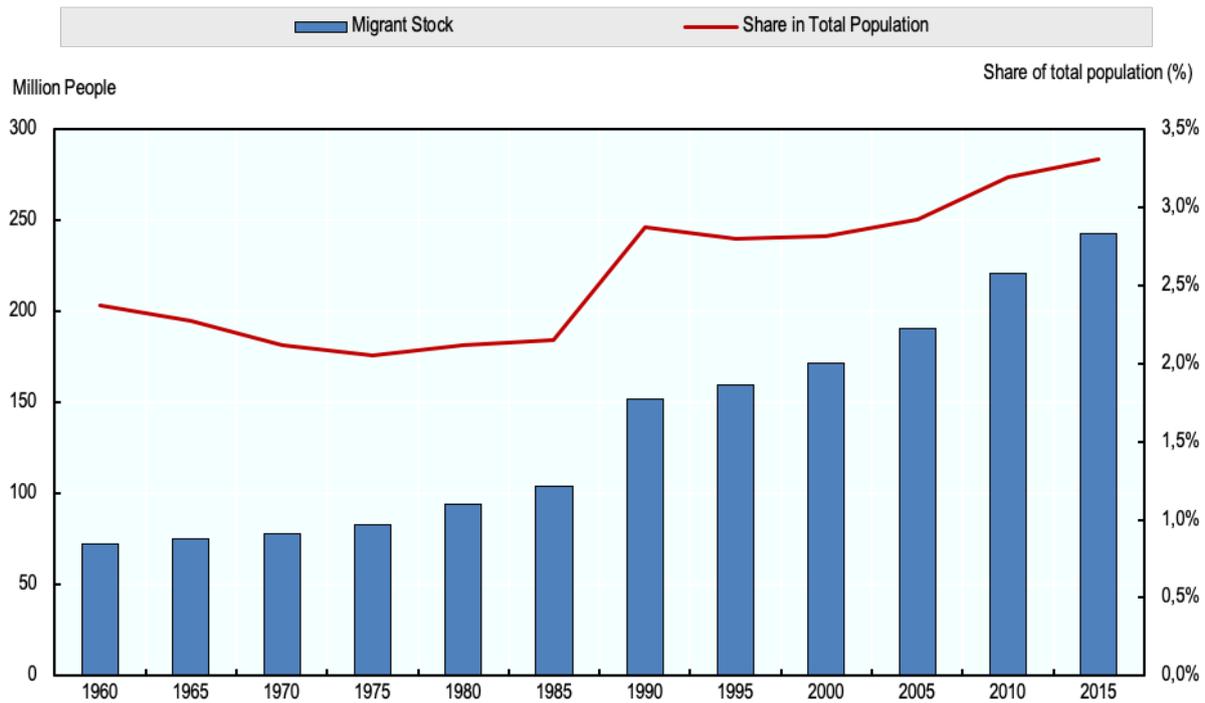
International Migration: What is all the fuss about?

Migration from places of economic disparity and the lack of opportunity is fundamental to our species existence. Dating back to the exodus of *Homo sapiens* out of eastern Africa, migration has been one of the great constants of human history. Long before political borders emerged, humans were walking the Earth, traveling in the hope of finding new and better lives. Some of these journeys were cyclical, such as the seasonal treks of nomadic tribes, others were more continuous journeys searching for a better home. Humanity's migration journey was not always voluntary; disruptions could take the form of wars, invasion, or famines. Whatever the reason may be, population movements across geographies created a long-lasting impact on the host and sending countries.

Until the early 20th century, migrants from much of the world traveled with few restrictions, provided they could afford and survive the administrative and health checks. At this time, immigrants rarely needed visas to make a new start in another country. Between 1920 and 1930, the trauma of unprecedented global warfare, economic recession, and growing xenophobia resulted in a sharp decline in the liberty of migration. By the time it re-emerged in the mid-1940s, the beginning of the "thirty glorious years," its character had changed.

Fast-forwarding today, around 243 million – or 3.3% – of the Earth's 7.3 billion people live outside their country of birth (2). Although this share in the global population remained roughly constant since 1960, the diversity in terms of composition (i.e., origin) and destination has increased. Today, international migration involves migrants originating from many more countries, going to many more destinations generating more origin-destination pairs than ever observed before (Özden et al., 2011). International migration remains a story of movements of people from less developed to more developed countries. Still, there is a substantial amount of movement between developed countries (i.e., "the north") and between developing countries ("the south"). Overall, about one-third of the world's migrants travel from north to north, another third travel from south to north, and the final third travel from south to south.

Figure 2: Foreign-born population



Note: Figure plots foreign-born stocks (left-axis) and as a share of total population (right-axis).
 Source: United Nations Migration Database

Just another factor of production in a globalized world?

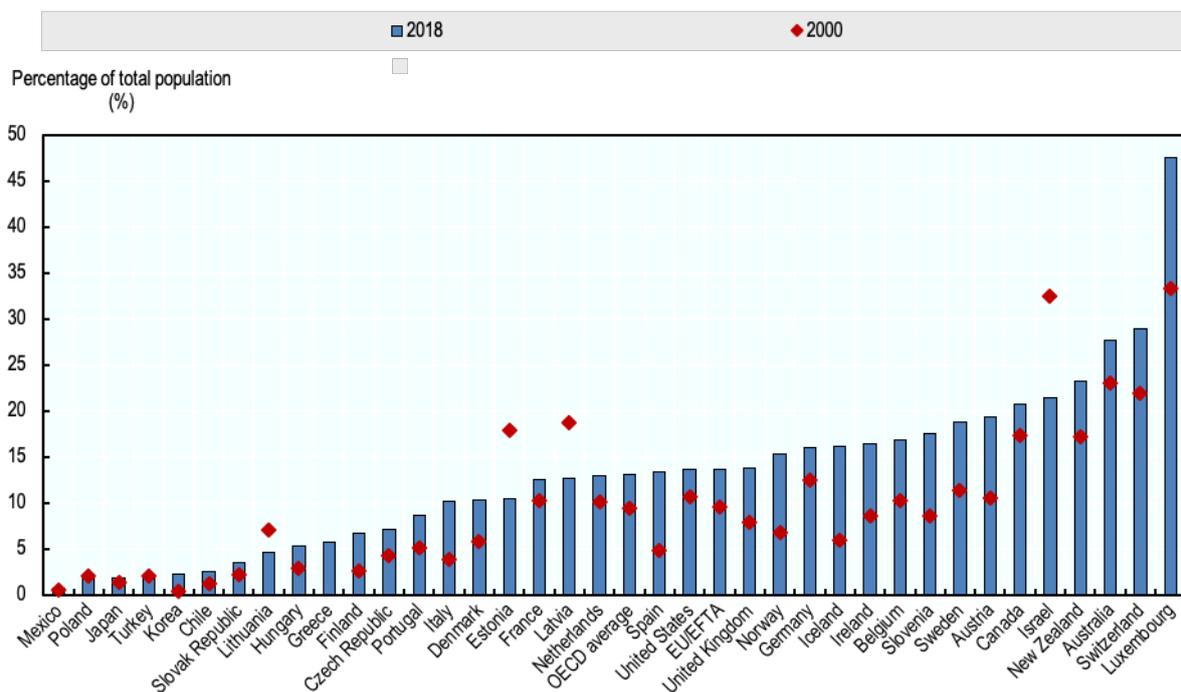
In the globalized world where capital, goods, and services move freely across borders, international migration remains limited. As more people look to live and work overseas, immigration has become a broad social and economic phenomenon. Although freedom of movement is not necessarily increasing, there is growing recognition of the role of migration as a component of globalization and a source of economic growth.

Despite evidence suggesting substantial efficiency and equity gains from freer international migration (Clemens, 2011), lowering barriers to immigration is not a popular idea supported in many dominating political circles. Going back in history, one will quickly realize that the cyclical nature of anti-immigrant attitudes is not limited to our current times. Today, the global news is occupied with US President Donald Trump’s constant thrum of anti-immigrant rhetoric and policy announcements and growing anti-immigrant attitudes in the US, Europe, and in war-bordering countries. As Howard Zinn shows in his popular book *People’s History of America*, resisting immigration has been a permanent part of American

history. As Zinn explains, members of the earlier waves of immigration always made the newcomers' lives difficult. Older immigrants from England and Germany were hostile toward the Irish, who had come during the Potato Famine between 1845 and 1855, and the Irish were subsequently hostile towards the Chinese immigrants 30 years later. Throughout the Age of Mass Migration (1850-1913) where about 30 million immigrants arrived in the United States, Italians, Jews, Polish, and Mexicans took their turn in being the subject of anti-immigrant rhetoric. Much like some immigrants today, these “new immigrants” were perceived by sectors of the US as being unlikely to assimilate.

Beyond the political debate, migration is a demographically important issue. For instance, between 2000 and 2018 (3), the increase in the foreign-born population accounted for more than three-quarters of the total population increase in European OECD countries, and for almost 40% of the rise in the United States (OECD, 2019). The increase in the demographic importance of international migration over the past decade has given more audience to the economic and political consequences of immigration.

Figure 3: Foreign-born population, by year and country



Note: Figure plots foreign-born stocks in 2000 (marker) and 2018 (bar).
 Source: OECD Migration Database

The recent increase in the interest in the migration question boosted the amount of research directed to the topic. As the issue gained public attention, researchers have tried to establish the impact of migration on the host and sending countries. For a long time, the literature focused on the impact of

immigration on labor-related outcomes of the native population. On the side of the sending country, most attention has been given to the importance of remittances for the economy and the cost of brain-drain on human capital accumulation. In the last decade, empirical literature expanded, providing a growing amount of evidence on migratory moves' impact on the economy and society in both their receiving and sending countries. Today, a growing number of papers show that migration boosts bilateral trade and investment, improve public finances, accelerate the diffusion of innovation, and generate shifts in values and political preferences. Furthermore, these effects can have heterogenous effects in the short and long-term.

Productivity and Migration

In addition to the benefits mentioned above, migrants also contribute to economic performance through productivity growth. Immigration can affect productivity through several channels: encouraging efficient specialization, diffusion of innovation, enhancing agglomeration externalities by increasing the density in urban areas, and skill-complementarities arising from diversity (see Peri (2016) for a summary).⁴

Migrants can also boost productivity by diffusing knowledge across firms, regions, and countries. For instance, as workers move between firms located within and across countries, they spread knowledge across space (Dasgupta, 2012; Poole, 2013b). As shown in Bahar and Rapoport (2018), migrants can play an important role as vehicles of knowledge transmission between the source and receiving countries, which can lead to increases in sectoral productivity and exports.

Migrants play a unique role in knowledge transmission due to their higher mobility rates compared to natives. Overall, immigrant workers more often move across regions, industries and occupations

⁴Diversity in “place of birth” of workers is argued to generate productive extenerlaties as it generates a greater variety of ideas, skills and goods and services supplied locally. The seminal reference is Ottaviano and Peri (2006), who find that birthplace diversity is positively correlated with both wages and rents in U.S. cities, which indicates diversity raises worker productivity. Similar studies in other advanced economies confirm this positive relationship (Bakens et al., 2013; Ager and Brückner, 2013; Alesina et al., 2016). These positive effects operate within the workplace as well as metropolitan scale (Trax et al., 2015; Kemeny and Cooke, 2018; Nathan, 2015).

compared to their native counterparts (Borjas, 2001; Orrenius and Zavodny, 2007). Migrant mobility, thus, can be a mechanism for reviving the diffusion machine across regions and country. By working in frontier firms located in the high productive regions, migrants are exposed to knowledge by the leading firms in their sector. At the end of migration episode, migrant returns back to his or her country of origin equipped with productive knowledge. Through this diffusion, countries can reduce interregional gaps and generate inclusive and balanced national growth. At the international level, countries lagging in terms of productivity and income can get a chance to improve productivity and catch-up to the leading economies.

Overall, this dissertation explores the consequences of mobility, both within and between countries, on productivity and labor markets. Across three empirical investigations in three different countries and period, I attempt to provide evidence on some aspects of migration that is little explored in the literature. Specifically, I explore the links between migrant mobility and labor markets' adjustment during a demand shock; migration as a channel for diffusing knowledge and boosting productivity across borders; and agglomerations and productivity gains associated with larger city sizes in developing economies.

1.1 Cushioning Effect of Immigrant Mobility: Evidence from the Great Recession

In the last two decades, the literature focused on the impact of immigration on the labor markets. In the first chapter, I reverse the mirror and explore an aspect little explored in the literature. I attempt to understand what happens to natives' labor market outcomes when migrants go home. Focusing on the context of Spain during the Great Recession, I study how the departure of immigrant workers impacts the wages and employment of native workers. Specifically, I attempt to show how the departure of foreign-born workers can dampen the effects of negative economic shocks on natives. To guide the empirical exercise and help the interpretation of my empirical findings, I commence by setting out a

theoretical framework. This simple model helps derive predictions about the native labor demand while accounting for both negative labor demand shock and a decrease in labor supply due to the departure of immigrant workers. The model's predictions show that a decrease in overall labor demand due to the economic crisis will decrease the labor demand for natives but also drive some immigrants away from the location. Faced with less competition from the immigrant workforce, labor market outcomes of natives improve. In the context of economic contraction and labor shedding, the decreased competition from immigrants will dampen the harmful effects of the crisis on native employment and wages. In reverse, in case of a positive demand shock, an increase in the labor demand for natives will be slowed by the rise in the competition due to the increased presence of immigrants from the same skill group. I test these predictions empirically using the context of Spain during the Great Recession. I focus on the direct impact of immigrants' outflow from local labor markets. I start by analyzing the changes in employment and wage outcomes of natives, located in each province, due to the reduction in labor supply due to the departure of immigrant workers. Specifically, I regress the annual growth in employment and wages of natives in a province on normalized net out-migration rates. To make causal claims about the estimates, I use a modified version of the standard shift-share instrument (Card, 2001). I analyze short-term effects on employment and wages for specific groups of workers (e.g., young or unskilled natives) but also types of employment adjustments in response to out-migration.

In my analysis, I use the Spanish Social Security panel, which allows me to track natives over time and across firms, sectors, and regions. I combine this data with individual-level municipal registers that cover 100% of the population to precisely count the number of immigrants residing in each province annually. I measure the intensity of departures across Spanish provinces by the annual decline in the immigrant working-age male population during this period, relative to the entire working-age male population in the previous year.

I find that the outflow of immigrants accelerated the wage and employment growth for natives of all skill and demographic groups. I find that the departure of immigrants increased geographic inflows of natives from other areas, increased the entry to employment from nonemployment, and improved the wages of those who were already employed.

Given the context of economic contraction, these findings show that through their higher mobility, immigrants diffuse the incidence of local shocks beyond the local labor market and cushion the natives during an adverse shock. In the context of wage stagnation and employment losses, the departure of immigrants, improved the wage growth while dampening the drop in employment. As the locations which were hit harder by the crisis also lost more immigrants, this smoothing effect worked like an automatic stabilizer stronger in hard-hit areas and worked as a channel for equilibrating differences across labor markets. This suggests that migrants do not only grease the wheels during the good times (Borjas, 2001) but also during bad times by leaving locations that are hit by a negative shock. This finding underlines an important benefit of immigration that has been explored little in the literature. This is particularly important given the concerns about the relative lack of mobility of natives and especially among less-skilled workers as it leads to significant divergence in local unemployment rates and workers' earnings across local labor markets (Bound and Holzer, 2002; Cadena and Kovak, 2016; Dao et al., 2017).

1.2 Migration and Post-conflict Reconstruction: The Effect of Returning Refugees on Export Performance in the Former Yugoslavia

The second chapter focuses on another novel angle exploring gains from migration: the role that migrants play in spreading ideas, technology, and knowledge across countries, and how such transfers are reflected in long-term real economic outcomes such as exports. With my co-authors Dany Bahar, Andreas Hauptmann, and Hillel Rapoport, we exploit a natural experiment and document how return migrants –having spent time in a foreign country– explain the subsequent performance of the same export sectors in which they had worked while abroad. Our study focuses on the early 1990s when about 700,000 citizens of the former Yugoslavia fled to Germany to escape the war. Most of the Yugoslavian migrants in the first half of the 1990s were given a temporary legal status, known as Duldung (German for “toleration”), which allowed them to stay and work in Germany. Duldung status was valid for six

months and was automatically renewed as long as the war was ongoing. In 1995 the Dayton peace agreements were signed, putting an end to the war in the former Yugoslavia. This led to Germany halting the renewing of the Duldung status of the Yugoslavian refugees, starting an enforced repatriation process. By 2000, the majority of Yugoslavian refugees had been repatriated back to their home country or to other territories of dissolved Yugoslavia.

Within this context, our study uses confidential administrative social security data from Germany that allow us to identify Yugoslavian refugees who had arrived in Germany during the Balkan refugee crisis and returned home after the war following 1995. In particular, we compute the number of these refugees who had worked in each one of the almost 800 four-digit tradable industries in Germany during that period (or ‘treatment’). We link this information to industry-level Yugoslavian export data. We estimate changes in export values from Yugoslavian countries to the rest of the world as a result of returning refugees who were employed in those same sectors in Germany.

We find that industries with a 10% higher number of returning workers have a larger level of exports to the rest of the world from 0.8% to 2.4% between the pre and post-war periods. Furthermore, these positive effects remain stable and keep increasing as time goes by. Our study rules out many explanations that could be driving the results, such as convergence, the inflow of investment linked to migration, or a decrease in information costs linked to international trade due to migrant networks.

Given that productivity is an underlying determinant of exports, we interpret these results as supporting the idea of migrants being drivers of know-how and technology transfers between countries. This interpretation is backed by the fact that our results are particularly driven by industries that are knowledge-intensive. They are also stronger when returnees are skilled and in occupations intensive in analytical and cognitive tasks.

A lesson to draw from our results is the importance of integrating refugees in their host economies’ labor markets, as this can really play a significant role in the post-conflict reconstruction of their home countries upon their return.

1.3 Agglomeration Effects in a Developing Economy: Evidence from Turkey

Urbanization has shaped developed countries during the 20th century, but it has had transformative effects on developing countries. Several major factors distinguish urbanization in developing countries from urbanization in developed countries. First, urbanization in developing countries is occurring at a quicker pace. In the last 50 years, on the back of a strong wave of rural-urban migration, emerging countries have seen the number of large cities increase. It took more than a century for most developed countries from the time urbanization started to increase markedly until they reached 50%. Today's developing countries often reach that threshold in less than half the time. Second, much larger numbers are involved making it a relevant issue for large numbers of people.

Given the importance of the topic, the third chapter of this dissertation focuses on understanding the determinants of regional productivity in a developing country context. For such a study, Turkey, an upper-middle-income developing country that has experienced fast urbanization and a high rate of growth of the urban population, seemed like a natural fit. Since the 1950s, the urban population in Turkey has increased dramatically due to massive rural-to-urban migration and a high fertility rate. In 1960, Turkey still featured a largely agrarian economy with 31 percent of its population residing in urban areas. By 2017, 75% of the Turkish population lived in cities, making it a very highly urbanized country. Today there are substantial inequalities between regions on almost every metric (income, production, life quality, etc.), including productivity.

In this study, I focus on understanding the sources of spatial productivity differences. For this purpose, I use social security records, a new administrative dataset that has only recently become available to researchers and thus has never been used in research before. Combining these records with various other datasets, I apply the standard two-step estimation approach to analyze the relationship between agglomeration economies and productivity in Turkey. Specifically, I attempt to measure the elasticity between employment density and productivity gains associated with denser urban areas. I address the endogeneity bias due to reverse causality by using historical instruments based on census data from the

Ottoman Empire and the early years of the Turkish Republic.

My results show that agglomeration economies are powerful and at levels that are expected given Turkey's urbanization level. The elasticity fits the literature perfectly and contributes to filling the knowledge gap about urbanization dynamics in the developing world. Beyond finding strong gains associated with denser agglomerations and market access, I also find that workers do not sort across locations based on their observable skills. This result is in sharp contrast with what is usually observed for developed countries, where a large fraction of the explanatory power of cities productivity advantages arises from sorting of the workers with higher abilities to denser and larger cities. Overall my findings corroborate earlier findings for developing countries, which show that while the main mechanisms of urban economies are present in the developing world, the current models need to be extended to capture the differences between Western cities and those in the developing world.

1.4 Concluding Remarks

Overall, the articles that constitute this dissertation aim to explore the interaction between migration and productivity, through multiple angles, across three different country and period contexts. Specifically, I study the labor market benefits of migrant mobility during an economic crisis, productivity gains due to migrant mobility in the reconstruction of a country in the aftermath of a war, and gains associated to the concentration of people coming from different parts of the economy, to generate higher growth and income. All of the chapters of this dissertation aim to understand the effects that arise from human mobility, across cities and countries. The findings of these studies relate to many issues that interest both the academia and the policymakers yet on which little is known. This dissertation aims to contribute to knowledge gap on issues that will remain relevant foreseeable future.

2 Introduction Generale

Qu'on les observe à l'âge reculé des chasseurs-cueilleurs, ou au sein de nos économies contemporaines, les migrations et la répartition spatiale de la productivité sont des phénomènes structurant de nos sociétés qui demeurent interconnectés. En un sens les migrations sont affectées par la répartition spatiale de la productivité. Si les hommes se déplacent c'est, dans bien des cas, 1) en quête d'opportunités apportées par la croissance économique, 2) dans l'espoir d'améliorer leur pouvoir d'achat ou 3) d'accumuler des richesses. Ces trois types de motivations dépendant fortement de la productivité, la répartition spatiale de la productivité joue un rôle significatif sur l'orientation des flux migratoires. Mais dans un autre sens, la répartition de la productivité est affectée par les migrations, car elle dépend non seulement des caractéristiques de l'individu, mais également du pays ou de la région dans laquelle il exerce son travail. Dès lors les flux migratoires doivent évoluer au grès des changements de productivité.

Durant la révolution industrielle, alors que les pays occidentaux surmontent les inerties caractéristiques des régimes de croissance pré-moderne, le reste du monde continue d'arborer des taux de croissance faible. S'opère alors une « Grande Divergence » (Pommeranz, 2000) des niveaux de productivité (et donc de revenus) entre les États occidentaux et le reste du monde qui, en s'accélégrant, s'accompagne de l'émergence d'inégalités territoriales au sein même des pays. De tels différentiels spatiaux engendrent alors mécaniquement des mouvements de population tant entre États qu'en leur sein même. Il n'est donc pas surprenant que les migrations demeurent aujourd'hui un enjeu de taille. Ces dernières années,

l'augmentation des flux migratoires dans le monde, ainsi que l'afflux de réfugiés à travers l'Europe, les placent au cœur des débats politiques qu'intellectuels.

A l'instar des phénomènes migratoires, les déterminants de la productivité font l'objet d'un intérêt accru depuis quelques années tant dans les milieux politiques qu'universitaires. Le ralentissement de la croissance depuis le milieu des années 2000, accentué au lendemain de la Grande Récession, est une cause évidente de ce regain d'intérêt. Améliorer le pouvoir d'achat par le truchement d'une croissance durable et inclusive constitue un incontournable de l'agenda politique des gouvernements du monde entier. Or la croissance de la productivité est une condition d'autant plus nécessaire à l'avènement d'une telle croissance, que le vieillissement démographique en menace la pérennité.

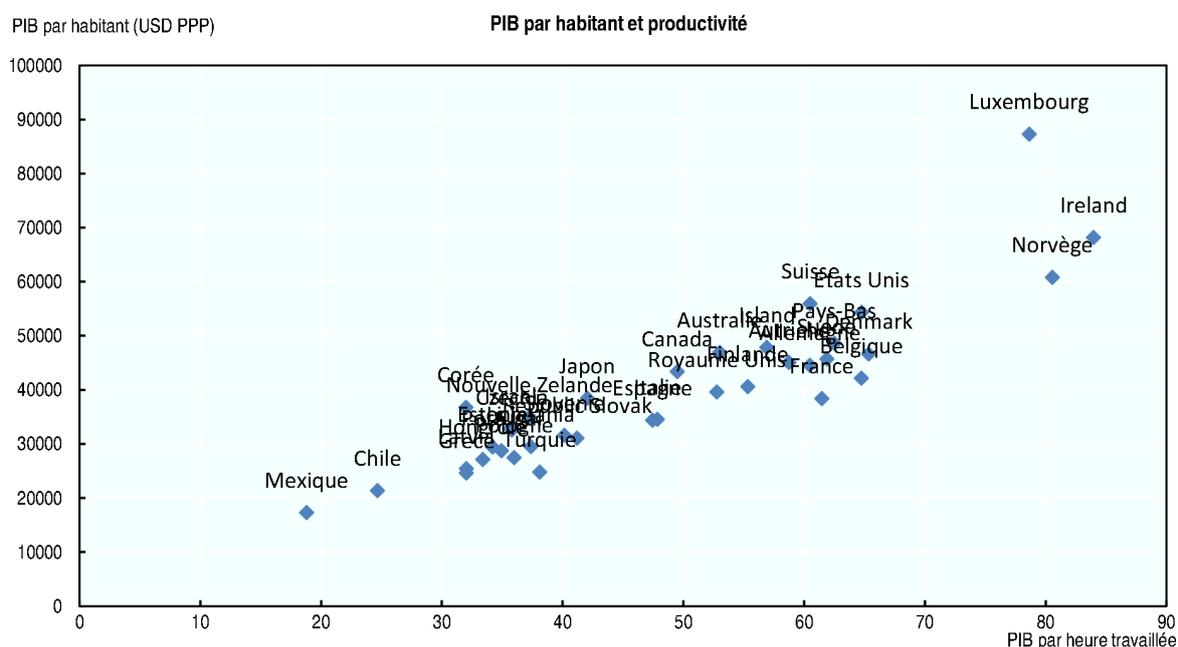
Ce regain d'intérêt s'impose également chez les universitaires, pour des raisons méthodologiques. L'usage des données migratoires apparaît comme de plus en plus à même de nous faire entrer dans cette « boîte noire » qu'est la productivité et, partant, d'identifier les canaux par lesquels elle peut être stimulée. Par conséquent, cette thèse est composée d'études qui se concentrent sur l'interaction entre phénomènes migratoires et la productivité, sous angles différents, en des espaces géographiques et des périodes divers.

Productivité : Pourquoi est-ce important?

La productivité d'un facteur est le rapport de la quantité d'output produite sur la quantité d'input nécessaire à la production de cet output. Pour le facteur de production qu'est le travail la productivité peut par exemple être donnée par la capacité d'output qu'un travailleur est capable de produire avec une quantité d'input donnée. Si un travailleur est plus productif qu'un autre c'est que pour la même quantité d'input le premier travailleur produit d'avantage d'output que le second travailleur. Un travailleur peut être rendu plus productif sous l'effet de nouvelles idées, d'innovations technologiques ou d'organisation du travail. Au début de la révolution industrielle, des innovations telles que la machine à vapeur, l'électrification et la numérisation permettent par exemple d'augmenter considérablement les rendements

de la production. Ils ont donc un effet positif sur la productivité. Etant donné la relation qui unie niveaux de productivité et salaires, ces derniers influencent les conditions matérielles, le niveau de vie et le bien-être de la population. En effet, les données historiques montrent que la croissance des salaires et de la productivité a progressé à un rythme soutenu au cours du dernier millénaire. Aujourd'hui, les grandes différences de revenu par habitant observées d'un pays à l'autre se reflètent principalement dans les différences de productivité du travail (4). La croissance de la productivité restera donc un moteur important de la croissance économique et de l'amélioration du bien-être.

Figure 4: Revenu par habitant et productivité du travail



Source: OECD

Le ralentissement de la croissance de la productivité depuis 2000

La croissance de la productivité dans la plupart des économies avancées a diminué depuis la fin des années 90. Malgré un pic temporaire consécutif à la crise, les taux de croissance baissent. Cela suscite des débats sur le caractère temporaire du ralentissement de la productivité (Brynjolfsson and McAfee, 2011) ou le signe d'un problème structurel plus permanent (Gordon, 2017).

La littérature scientifique offre diverses explications au ralentissement de la productivité.⁵ L'une des

⁵Pour de plus amples informations sur la nature et les causes du ralentissement de la productivité, voir, entre autres

explications avancées est la panne de ce qu'on appelle la "machine à diffusion". Selon cet argument, bien que les grandes entreprises (ou celles qui sont à la "frontière" technologique) continuent d'innover et d'accroître leur productivité, les connaissances productives générées dans ces entreprises ne se diffusent plus aux entreprises moins productives. Un tel phénomène génère des différences de productivité au sein des pays et des industries.

Bien qu'il reste encore à trouver une explication claire à ce problème de diffusion, il est amplement prouvé que la géographie joue un rôle important. Dans de nombreux pays de l'OCDE, la croissance de la productivité observée depuis les années 2000 n'est générée que par quelques régions, tandis que les autres stagnent, quand elles ne souffrent pas d'une baisse de productivité. En conséquence, de nombreux pays de l'OCDE connaissent des augmentations substantielles de la productivité interrégionale et des disparités économiques.⁶ Alors que les écarts de PIB par habitant entre les pays de l'OCDE se réduisent depuis la fin des années 90, les pays doivent faire face à des écarts de revenu croissants dans leurs propres régions et villes.

Ces disparités persistantes entre les régions alimentent un processus de divergence intra-étatique en matière économique et sociale. Les mauvais résultats économiques, le taux de chômage toujours élevé et la baisse des salaires diminuent les perspectives et les capacités de certaines collectivités locales à avoir un avenir meilleur. Par conséquent, dans de nombreux pays, une part croissante de la population est de plus en plus mécontente du statu quo. Il n'est pas surprenant que ce mécontentement ait un profil géographique et qu'il soit fortement corrélé avec les performances économiques des régions (McCann 2019; Rodríguez-Pose 2018).

Villes : Source de croissance de la productivité

Les villes sont un moteur clé de la performance économique et de la croissance de la productivité de leurs régions. Dans l'économie numérique et de services d'aujourd'hui, les villes facilitent des interactions

(Cette, Corde, and Lecat 2018; Syverson 2017).

⁶Les chocs commerciaux et la concurrence accrue des pays en développement dans le secteur manufacturier sont une autre explication avancée pour expliquer la divergence des niveaux de productivité et de revenu entre les régions de certains pays développés. Pour une revue de la littérature sur les effets de la concurrence commerciale chinoise sur le marché du travail, voir par exemple (Autor, Dorn, and Hanson 2016).

génératrices de croissance entre industries transformatrices et secteurs à haute intensité de connaissance.¹ En facilitant la mobilité et l'interaction entre travailleurs, la proximité des agents économiques favorise la diffusion des savoirs et génère d'importantes retombées technologiques. Bien que l'innovation puisse, en principe, se produire n'importe où, on l'observe principalement dans les zones fortement urbanisées en raison de la forte dégradation spatiale des connaissances (Carlino and Kerr, 2015). La croissance de la productivité générée dans les villes se répercute également sur les régions environnantes et le reste de l'économie, ce qui influe sur leur performance économique. Ainsi, en repoussant la frontière de la productivité et en générant une croissance de la productivité, les villes deviennent des moteurs essentiels du potentiel de croissance économique à long terme de leur région et de du pays.

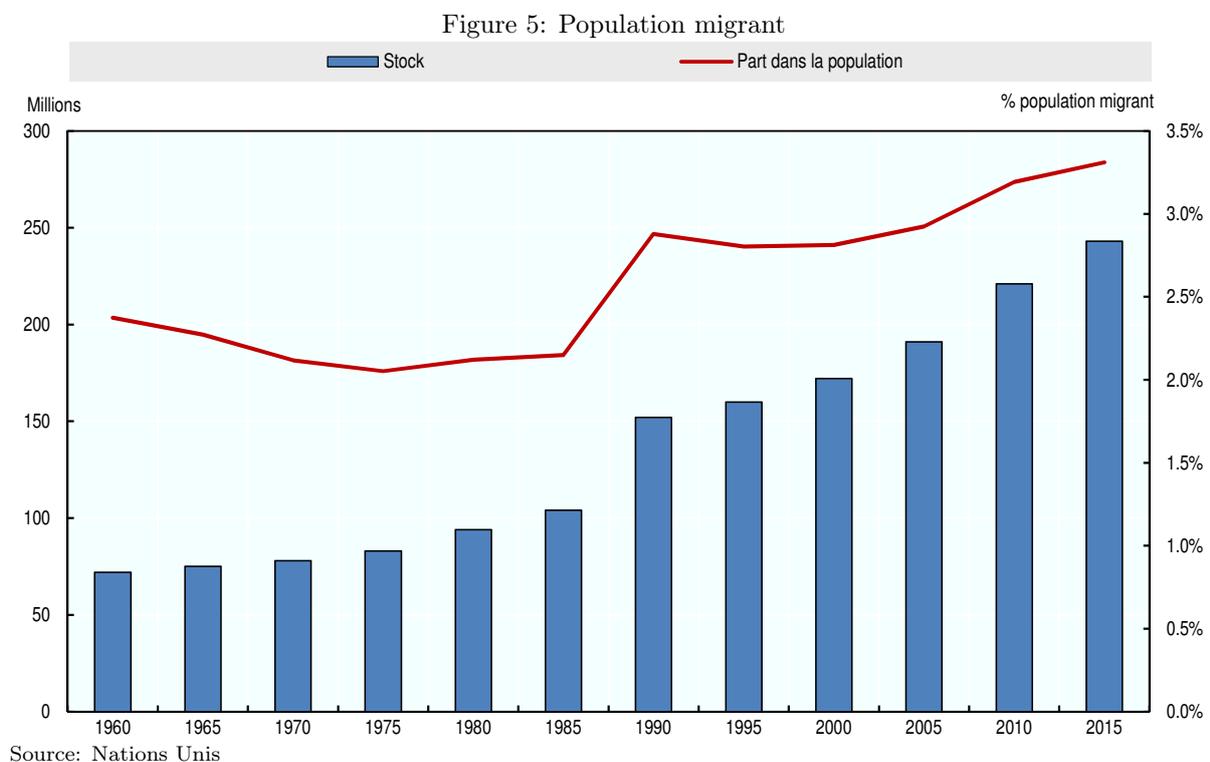
L'urbanisation dans les pays en développement joue un rôle important dans la transformation structurelle et la croissance économique des pays émergents. Les villes sont devenues des pôles économiques, pionnières du reste de l'économie non seulement en matière d'innovation et de productivité, mais en matière de politiques culturelles, éducatives et sanitaires. Le puissant pôle d'attraction économique, associé à des commodités attrayantes, continue d'attirer l'exode rural, qui est le principal moteur de la croissance démographique urbaine dans de nombreux pays en développement. En conséquence, bien que ces villes jouissent de niveaux de productivité élevés et parfois d'une qualité de vie élevée, la plupart d'entre elles sont victimes de leur succès. Dans de nombreux pays en développement, l'absence d'infrastructures institutionnelles et physiques adéquates pour absorber un afflux important de populations augmente les coûts d'agglomération, réduisant ainsi les gains consécutifs à la densification et à l'extension.

Comprendre l'importance des agglomérations dans les pays en développement est précieux car, au-delà de leur taille et de leur poids dans l'économie globale, elles sont au cœur du changement structurel et restent la première destination des migrants internes qui cherchent une vie meilleure.

International Migration : De quoi s'agit-il ?

L'exode des premiers hommes hors d'Afrique de l'Est est le premier exemple de migration. Depuis lors, la migration est une constante de l'histoire humaine. Bien avant l'apparition des frontières politiques, les hommes parcouraient la Terre, voyageant dans l'espoir de trouver de nouvelles ressources et de meilleures conditions d'existence. Certains de ces voyages étaient cycliques, comme les randonnées saisonnières des tribus nomades, d'autres étaient des voyages plus continus à la recherche d'un meilleur endroit pour vivre. Les migrations ne sont pas toujours volontaires ; elles peuvent être la conséquence de guerres, d'invasions, de famines... Quelle qu'en soit la raison, ces mouvements de population d'une région ont un impact durable tant sur le pays d'accueil que sur le pays d'origine.

Jusqu'au début du XXe siècle, les migrations sont peu réglementées. Les migrants doivent simplement passer les contrôles administratifs et sanitaires, et ont rarement besoin de visas pour immigrer. Entre 1920 et 1930, le traumatisme d'une guerre mondiale sans précédent, la récession économique et la xénophobie croissante va de paire avec la mise en place de réglementations plus restrictives. Au milieu des années 1940, au début des "trente glorieuses années", son caractère avait changé.



En ce début de XXI^e siècle, les migrations internationales sont toujours entravées par des restrictions juridiques. Ces dernières années, elles s'imposent même comme un thème dominant et comme une source inaltérable de controverses économiques et politiques. En 2015, 243 millions - soit 3,3 % - des 7,3 milliards des hommes qui peuplent le monde vivent hors de leur pays de naissance (5). Bien que cette part de la population mondiale soit à peu près constante depuis 1960, la diversité en termes de composition (c.-à-d. d'origine) et de destination augmente. Aujourd'hui, les migrations internationales se composent des migrants originaires de beaucoup plus de pays, qui se rendent vers des destinations plus variées, générant plus de couples origine-destination que jamais observés auparavant (Özden et al., 2011).

Les phénomènes migratoires n'en restent pas moins dirigés depuis des pays moins développés vers les pays qui le sont davantage. Cela ne doit pas cependant cacher les mouvements entre les pays développés (c.-à-d. le " Nord ") et entre les pays en développement (le " Sud "). Dans l'ensemble, environ un tiers des migrants du monde voyagent du Nord vers le Nord, un autre tiers du Sud au Nord et le dernier tiers du sud au sud.

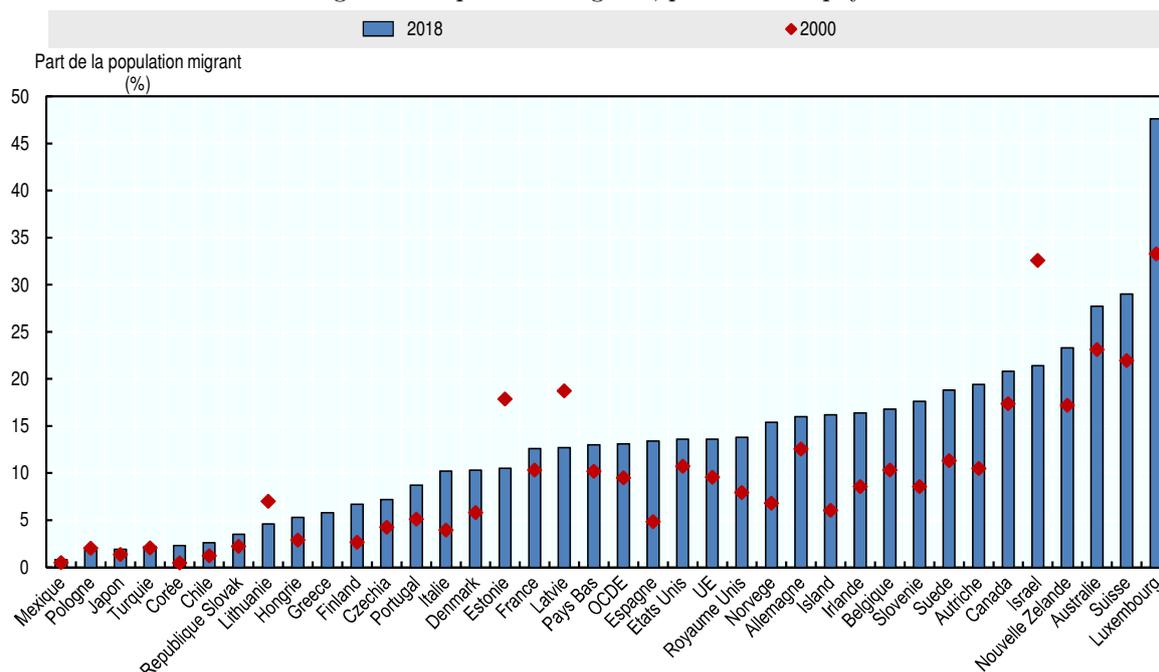
Juste un autre facteur de production dans un monde globalisé?

Dans notre monde globalisé les migrations internationales restent limitées, en comparaison des mouvements capitaux, des biens et les services dont la circulation est de moins en moins entravée aux frontières. Comme de plus en plus de gens cherchent à vivre et à travailler à l'étranger, l'immigration est devenue un vaste phénomène social et économique. Bien que les réglementations ne s'assouplissent guère, le rôle de la migration comme composante de la mondialisation et source de croissance économique est de plus en plus reconnu.

Malgré les preuves suggérant des gains substantiels d'efficacité et d'équité découlant d'une libéralisation des flux migratoires (Clemens, 2011), réduire les obstacles à l'immigration est loin d'être une idée populaire. Son rejet n'est pas cependant l'apanage de notre époque. Il suffit de remonter dans l'histoire, pour apprécier le caractère cyclique des attitudes anti-immigrées. L'actualité mondiale est par exemple inondée de la vindicte anti-immigrés du président américain Donald Trump. Cela laisse augurer la mul-

tiplication de politiques, de discours et d'attitudes hostiles à l'égard de l'immigration aux États-Unis, en Europe et dans les pays frontaliers de la guerre. Comme le montre Howard Zinn dans son livre populaire *People's History of America*, la résistance à l'immigration fait partie intégrante de l'histoire américaine. Comme l'explique Zinn, les membres des premières vagues d'immigration ont toujours rendu la vie des nouveaux arrivants difficile. Ainsi les immigrants d'Angleterre et d'Allemagne, issues d'une génération antérieure, se montrent hostiles à l'égard des migrants Irlandais débarqués pendant la famine de 1845 à 1855. Ainsi les Irlandais font-ils preuve d'agressivité envers les immigrants chinois 30 ans plus tard. Tout au long de l'ère des migrations de masse (1850-1913), où environ 30 millions de personnes immigreront aux États-Unis, les Italiens, les Juifs, les Polonais et les Mexicains ont à leur tour fait l'objet de rhétorique anti-immigrants. Tout comme certains immigrants aujourd'hui, ces " nouveaux immigrants " sont alors perçus dans certaines régions des États-Unis comme inassimilable.

Figure 6: Population migrant, par année et pays



Source: OECD Migration Database

Au-delà du débat politique, les migrations constituent un enjeu démographique important. Par exemple, entre 2000 et 2018 (6), l'augmentation de la population née à l'étranger représente plus des trois quarts de l'accroissement total de la population dans les pays européens de l'OCDE et près de 40 % de l'augmentation aux États-Unis (OECD Migration Outlook, 2019). L'intensification des flux migra-

toires au cours de la dernière décennie permet de mieux faire connaître les conséquences économiques et politiques de l'immigration.

Bien qu'il y ait un intérêt de longue date pour l'étude des migrations dans les pays d'accueil, dans sa majeure partie la littérature se concentre sur son impact sur le domaine du travail, sur le niveau des salaires et le niveau d'emploi principalement. Du côté des pays émetteurs, l'attention est portée sur l'impact des transferts de fonds et de la fuite des cerveaux. De plus en plus d'éléments montrent que les migrations ont un impact sur de nombreux aspects de l'économie et de la société, tant dans les pays d'accueil que dans les pays d'origine : elles stimulent le commerce et les investissements bilatéraux, augmentent l'assiette fiscale, accélèrent la diffusion de l'innovation et entraînent des changements dans les valeurs et les préférences politiques. De plus, ces effets peuvent se produire à court et à long terme.

Les gains de productivité dus à l'agglomération des personnes et des entreprises ne se limitent pas aux retombées du savoir. L'économiste Alfred Marshall (Marshall 1890) a été parmi les premiers à souligner que l'agglomération de personnes et d'entreprises peut accroître la productivité. Depuis lors, des efforts considérables ont été consacrés à l'identification des mécanismes par lesquels les économies d'agglomération émergent. Bien que tous les aspects liés à l'émergence des économies d'agglomération ne soient pas bien compris, les avantages peuvent être répartis en trois grands groupes : partage, appariement et apprentissage (Duranton et Puga, 2004).

Productivité et migration

En faisant croître la productivité, les migrants contribuent aux performances économiques des États. L'immigration affecte la productivité de plusieurs manières : elle encourage une spécialisation efficace, elle diffuse l'innovation, elle renforce les externalités d'agglomération en augmentant la densité dans les zones urbaines et les complémentarités de compétences résultant de la diversité (voir Peri (2016) pour une revue de la littérature).

La diversité dans le "lieu de naissance" des travailleurs est considérée comme génératrice d'extensibilités productives car elle génère une plus grande variété d'idées, de compétences et de biens et services fournis localement. La référence fondamentale est Ottaviano et Peri (2006), qui constatent que la diversité des lieux de naissance est corrélée positivement aux salaires et aux loyers dans les villes américaines, ce qui indique que la diversité augmente la productivité des travailleurs. Des études similaires dans d'autres économies avancées confirment cette relation positive (Nathan, 2011; Kemeny, 2012; Bakens et al., 2013; Ager et Brückner, 2013; Suedekum et al. 2014 ; Elias et Paradies, 2016; Alesina et al. 2016). Ces effets positifs se manifestent tant à l'échelle du milieu de travail qu'à l'échelle métropolitaine (Trax et al., 2015; Kemeny et Cooke, 2015; Nathan, 2016).

Bien que cela soit peu documenté, les migrations sont également facteurs de productivité en ce qu'elles favorisent la diffusion des connaissances à travers les entreprises, les régions et les pays. Comme nous l'avons vu dans la littérature sur les retombées de la productivité dues à la diffusion des connaissances dans les entreprises par le biais de la mobilité des travailleurs d'une entreprise à l'autre (Dasgupta, 2010; Poole, 2013), les migrants peuvent aussi être un canal de diffusion des connaissances entre les régions et les pays. Comme le montrent Bahar et Rapoport (2018), les migrants peuvent jouer un rôle important en tant que vecteurs de connaissances entre les pays d'origine et les pays d'accueil, ce qui conduit à une productivité sectorielle et à des exportations accrues.

Les immigrants sont plus mobiles que ce soit en matière géographique, ils sont plus enclin à changer de région pour trouver du travail, ou que ce soit sur le marché du travail, ils acceptent davantage de changer de secteur d'activité (Borjas, 2001 ; Orrenius et Zavodny, 2007). La mobilité des migrants peut donc être un mécanisme permettant de relancer la machine de diffusion à travers les régions et les pays. Grâce à leur grande mobilité par rapport aux natifs, ils sont porteurs de connaissances entre entreprises, régions et pays. En travaillant dans des entreprises frontalières situées dans les régions les plus performantes, les migrants sont exposés aux connaissances des principales entreprises de leur secteur. A la fin de l'épisode migratoire, le migrant rentre dans son pays d'origine avec des connaissances productives. Grâce à cette diffusion, les pays peuvent réduire les écarts interrégionaux et générer une croissance nationale inclusive et équilibrée. Au niveau international, les pays en retard en termes de productivité et de revenus peuvent

avoir une chance d'améliorer leur productivité et de rattraper leur retard par rapport aux principales économies.

Cette thèse explore le lien entre la productivité et les migrations internationales à travers trois études empiriques et étudie l'interaction des deux facteurs sous des angles très divers. Plus précisément, Nous interrogeons les liens entre la mobilité des migrants et l'ajustement des marchés du travail en période de choc de la demande, la migration comme moyen de diffuser les connaissances et d'accroître la productivité au-delà des frontières, et les agglomérations et les gains de productivité associés aux villes de grande taille.

2.1 Effet Amortisseur de la Mobilité des Immigrants

Historiquement, la littérature économique s'est d'abord attachée à étudier l'impact de l'immigration sur le marchés du travail. Dans le premier chapitre de cette thèse, nous inversons la focale, et nous intéressons à un aspect peu documenté : l'effet du retour des migrants dans leur pays d'origine sur l'emploi et le salaire des natifs. Plus précisément, nous démontrons comment le départ de travailleurs nés en dehors du territoire espagnol après la crise de 2008 a potentiellement atténué les effets négatifs de cette crise sur l'emploi et les salaires des natifs.

Pour guider l'exercice empirique et faciliter l'interprétation de ses résultats, nous nous sommes employé à construire un modèle simple. Ce cadre théorique permettra de dériver des prédictions qui résultent d'un double effet sur la demande de travailleurs natifs: un effet négatif de la crise sur la demande globale de travail, un second effet positif lié à la diminution de l'offre de travail consécutive au départ des travailleurs migrants. Ces prévisions suggèrent donc, en ce qui concerne les natifs, qu'un effet négatif de la crise sur le marché travail pourrait être, au moins partiellement, compensé par le départ des migrants. Confrontée à une concurrence moindre, la situation des natifs sur le marché du travail, dès lors, s'améliore. A l'inverse, l'effet d'un choc positif sur la demande de travailleurs natifs, est atténuée

par l'arrivée de nouveaux migrants appartenant au même groupe de compétence, et par extension, par l'intensification de la concurrence.

Pour tester empiriquement ce cadre théorique, nous nous appuyons sur des données espagnoles pendant la Grande Récession. L'analyse se concentrera sur l'impact direct de l'exode des immigrants sur le marché du travail au sein des différentes provinces espagnoles. Nous avons analysé dans un premier temps, l'évolution de l'emploi et des salaires des natifs, pour chacune de ces provinces. Plus précisément, je régresse la croissance annuelle de l'emploi et des salaires des natifs d'une province sur les taux nets normalisés de migration de sortie. Pour faire des affirmations causales au sujet des estimations, nous avons employé une version modifiée de l'instrument standard de partage des quarts de travail (Card, 2001). Sur cette base, nous a permis d'étudier les effets à court terme sur l'emploi et les salaires de groupes particuliers de travailleurs (p. ex., les jeunes natifs ou les travailleurs non qualifiés), mais aussi les façons dont s'ajuste l'emploi en réponse à l'émigration.

Dans cette analyse, les données de panel sont issues de la sécurité sociale espagnole, permettant un suivi des natifs dans le temps et entre entreprises, secteurs et régions. Nous avons combiné ces données avec les registres municipaux individuels qui couvrent 100 % de la population pour compter précisément le nombre d'immigrants résidant dans chaque province chaque année. Nous mesurons ensuite l'intensité des départs dans les provinces espagnoles par la diminution annuelle de la population masculine immigrante en âge de travailler au cours de cette période, comparativement à l'ensemble de la population masculine en âge de travailler l'année précédente.

Force est de constater que l'exode des immigrants a accéléré la croissance des salaires et de l'emploi pour les natifs pour tous les groupes de compétences et pour tous les groupes démographiques. Le départ des migrants a ainsi eu un effet positif sur la mobilité interprovinciales des natives et, en conséquence, a permis d'augmenter l'entrée sur le marché du travail à partir du non-emploi et amélioré les salaires de ceux qui en possédaient déjà un.

Compte tenu du contexte de contraction économique, ces résultats suggèrent qu'en raison de leur plus grande mobilité, les immigrants participent à une diffusion des chocs locaux au-delà du marché du travail local, et, en cas de choc négatif, en amortissent l'incidence. Dans un contexte de stagnation des salaires

et de pertes d'emploi, le départ des immigrants a un effet positif sur la croissance des salaires tout en freinant la baisse de l'emploi. Étant donné que les régions les plus durement touchées par la crise aussi été celles ayant perdus le plus d'immigrants, cet effet de lissage a fonctionné comme un stabilisateur automatique plus puissant dans les régions durement touchées et a servi de canal pour équilibrer les différences sur les marchés du travail. Cela suggère que les migrants ne font pas que « graisser les roues pendant les bons moments » (Borjas, 2001), mais aussi pendant les mauvais moments en quittant des endroits qui sont touchés par un choc négatif. Cette constatation souligne un avantage important de l'immigration qui reste peu documenté. C'est d'autant plus important que le manque de mobilité des natifs, en particulier des moins qualifiés, entraînent des déséquilibres territoriaux significatifs, que ce soit en matière d'emploi, qu'en matière de salaire (Bound et Holzer, 2002; Cadena et Kovak, 2016; Dao et al., 2017).

2.2 Migration et Reconstruction Post-conflit : L'Effet du Retour des Réfugiés sur les Exportations en ex-Yougoslavie

Le deuxième chapitre s'intéresse à un des gains potentiels des migrations : le rôle que jouent les migrants dans la diffusion des idées, des technologies et des connaissances entre les pays, et comment ces transferts se reflètent dans les résultats économiques réels à long terme, comme les exportations. Dany Bahar, Andreas Hauptmann, Hillel Rapoport et moi-même, exploitons une expérience naturelle : l'exil forcé des populations d'ex-Yougoslavie consécutive aux guerres des Balkans et leur retour plusieurs années après le conflit. Nous montrons comment leur séjour à l'étranger explique les performances ultérieures à l'export, dans les secteurs au sein desquels ils avaient travaillé à l'étranger.

Notre étude porte sur le début des années 1990, lorsque quelques 700 000 citoyens de l'ex-Yougoslavie fuient vers l'Allemagne pour échapper à la guerre. A l'époque, la plupart d'entre-eux obtiennent un statut juridique temporaire, connu sous le nom de Duldung (tolérance en Allemand), qui leur permet de

rester et de travailler sur le territoire. D'une durée de validité de six mois, ce statut est automatiquement renouvelé tant que la guerre perdure. En 1995, les accords de paix de Dayton sont signés, mettant fin à la guerre en ex-Yougoslavie. Le statut de Duldung cesse alors d'être renouvelable forçant les immigrants yougoslaves au retour. En 2000, la majorité des réfugiés d'origines yougoslaves sont de retour dans leur pays d'origine ou dans d'autres territoires de la Yougoslavie dissoute.

Dans ce contexte, notre étude utilise des données administratives exclusives et confidentielles de la sécurité sociale allemande nous permettant d'identifier les réfugiés yougoslaves arrivés en Allemagne pendant la crise des Balkans et de retour chez eux après la guerre à partir de 1995. Nous calculons en particulier le nombre de ces réfugiés qui ont travaillé dans chacune des quelque 800 industries échangeables à quatre chiffres en Allemagne pendant cette période (dite de " traitement "). Nous relierons par la suite ces informations aux données sur les exportations industrielles yougoslaves et estimons les variations de la valeur des exportations des pays yougoslaves vers le reste du monde à la suite du retour des réfugiés qui travaillaient dans ces mêmes secteurs en Allemagne.

Dans notre étude, nous constatons que les industries qui comptent 10 % ou plus de travailleurs de retour ont un niveau plus élevé d'exportations vers le reste du monde, situés entre 0,8 % et 2,4 % entre les périodes d'avant et d'après-guerre. De plus, ces effets positifs restent stables et continuent d'augmenter avec le temps. Notre étude exclut de nombreuses explications, telles que la convergence, l'afflux d'investissements liés à la migration, ou la diminution des coûts d'information liés au commerce international en raison des réseaux de migrants.

Étant donné que la productivité est un déterminant sous-jacent des exportations, nous interprétons ces résultats comme appuyant l'idée que les migrants sont les moteurs des transferts de compétences et technologiques entre pays. Cette interprétation est renforcée par le fait que nos résultats s'expliquent principalement par l'impact des industries à forte intensité de connaissances. Ils sont également plus forts lorsque les rapatriés sont qualifiés et occupent des emplois intensifs dans des tâches analytiques et cognitives

2.3 Economies d'Agglomération dans un Economie en Développement : Le Cas de la Turquie

Au cours du XXe siècle, si le processus d'urbanisation n'épargne que très peu d'Etats du monde, son déroulement effectif diffère entre pays développés et pays en développement, et ce pour au moins deux raisons. Premièrement, l'urbanisation dans les pays en développement se fait à un rythme beaucoup plus soutenu que dans les pays développés. Alors qu'il faut tout un siècle aux pays développés pour atteindre les 50 %, aujourd'hui, la plupart des pays émergents n'atteignent ce seuil qu'en un demi siècle. Deuxièmement, les ordres de grandeurs sont beaucoup plus importants dans les pays en développement : l'urbanisation dans les pays en développement concerne donc un grand nombre de personnes. Les conséquences d'un tel phénomène revêtent donc une importance considérable, notamment sur la répartition spatiale de la productivité au sein de ces Etats.

Pour cette raison, nous nous concentrons dans ce troisième chapitre de sur la compréhension des déterminants de la productivité régionale dans le contexte d'un pays en développement : la Turquie. A partir des années 1950, la population urbaine turque augmente de façon spectaculaire en raison de l'exode rural massif et d'un taux de fécondité élevé. Alors qu'en 1960, elle présente encore une économie essentiellement agraire avec seulement 31 % de citadins, en 2017, 75% de la population turque vit dans les villes, ce qui en fait un pays très fortement urbanisé. Aujourd'hui, il existe d'importantes inégalités entre les régions sur presque tous les paramètres (revenu, production, qualité de vie, etc.), y compris la productivité.

Dans cette étude, nous insistons sur la compréhension des sources des différences de productivité spatiale. Pour ce faire, nous avons utilisés les dossiers de sécurité sociale, un nouvel ensemble de données administratives mis à la disposition des chercheurs récemment et qui n'a donc jamais été utilisé dans un cadre scientifique auparavant. En combinant ces enregistrements avec divers autres ensembles de données, nous appliquons l'approche standard d'estimation en deux étapes pour analyser la relation entre les économies d'agglomération et la productivité en Turquie et plus précisément, en essayant de

mesurer l'élasticité entre la densité d'emploi et les gains de productivité associés aux zones urbaines plus denses. Une importance particulière sera accordée au biais d'endogénéité dû à la causalité inverse en utilisant des instruments historiques basés sur les données de recensement de l'Empire ottoman et des premières années de la République turque.

Les résultats obtenus démontrent l'importance des économies d'agglomération et compte tenu du niveau d'urbanisation de la Turquie. L'élasticité correspond parfaitement à la littérature et contribue à combler le fossé des connaissances sur la dynamique de l'urbanisation dans les pays en développement. Au-delà des gains importants associés aux agglomérations plus denses et à l'accès au marché, nous constatons également que les travailleurs ne trient pas les données en fonction de leurs compétences observables. Ce résultat contraste fortement avec ce qui est généralement observé dans les pays développés, où une grande partie du pouvoir explicatif des avantages des villes en termes de productivité provient du tri des travailleurs les plus aptes à travailler dans des villes plus denses et plus grandes. Dans l'ensemble, les résultats obtenus corroborent les constatations antérieures pour les pays en développement qui démontrent que si les principaux mécanismes des économies urbaines sont présents dans le monde en développement, les modèles actuels doivent être élargis pour saisir les différences entre les villes occidentales et celles du monde en développement.

3 The Cushioning Effect of Immigrant Mobility: Evidence from the Great Recession in Spain

Abstract

This paper provides the first direct evidence on how the labor mobility of immigrants cushions natives during a labor demand shock. Spain was one of the hardest-hit economies during the Great Recession. Faced with a drop in the local labor demand, immigrant workers moved to other locations in Spain or left the country. Focusing on this episode, using microdata from municipal registers and longitudinal Spanish administrative data, I study the effects of out-migration of the immigrant population on the wages and employment of the remaining natives. I build a shift-share instrument based on the past settlements of the immigrant population across Spain to instrument outflows and argue for a causal relationship. I find that out-migration of immigrants accelerated employment and wage growth of the natives, especially for those with higher substitutability with the leaving population. Moreover, I find that employment effects are driven by increased entry to the employment of individuals who were unemployed or inactive, while wage effects were limited to those who were already employed. These findings indicate that through their mobility, immigrants diffused the incidence of local shocks and cushioned the fall of the natives.

Keywords: emigration, wages, employment, local labour market

JEL Classification Numbers: F22, J31, J61, R23

3.1 Introduction

Despite an extensive literature on the effects of the arrival of immigrants on the labor market outcomes of the native population,⁷ the literature on the impact of their departure remains limited. The absence of literature is surprising as mobility of the immigrant population can be an important channel for equilibrating labor markets (Cadena and Kovak 2016; Basso et al. 2018; Jauer et al. 2019). Due to the higher responsiveness to labor market differentials, immigrants react to the changes in labor market conditions through mobility much more than natives (Borjas and Aydemir, 2007; Schündeln, 2014). While natives in a distressed local markets stay in the area, immigrants relocate to other markets and dissipate the shock spatially (Monras, 2018). Through their mobility, immigrants can absorb part of the shock and help the local labor market recover faster.

Spain was one of the hardest hit economies during the Great Recession. The immigrant population in the country responded to the decreased local labor demand by moving internally but also leaving the country. Between 2009-2016, Spain saw a net reduction of 300 000 of its immigrant working-age male population.⁸ The outflow of the immigrant male population resulted in a decrease of 2% of the country's total male labor supply.⁹

In this paper, I examine the direct impact of immigrant out-migration on the local labor markets over the 2009-2014 period.¹⁰ I start by analyzing the changes in employment and wage outcomes of natives located in each province due to labor supply reduction generated by the departure of immigrant workers. Specifically, I regress the annual growth in employment and wages of natives in a province on normalized net out-migration rates. To make causal claims about the estimates, I use a modified version of the standard shift-share instrument (Card, 2001). I analyze short-term effects on employment and wages for specific groups of workers (e.g., young or unskilled natives) but also types of employment adjustments in response to out-migration. For example, although native employment adjustments can be through higher inflows from non-employment into employment, it can also result from fewer employed workers leaving the labor force. Furthermore, the adjustment can also be through geographic movements across local labor markets, a mechanism found to be essential in explaining the long-run effects of adverse demand shocks in the United States (see Blanchard and Katz, 1992). I provide evidence on the magnitude of each type of response and show how their relative importance varies across worker groups.

To guide my empirical investigation, I start with a general equilibrium model that predicts how a decrease in the local labor supply due to immigrant out-migration affects wages and employment of the

⁷See for instance Altonji and Card (1991); Card (2001); Angrist and Kugler (2003); Glitz (2012); Ottaviano and Peri (2012); Dustmann et al. (2017a).

⁸At the end of 2014 the number of immigrant citizens exceeded the number of population with immigrant nationality by over one million and a half. For this reason, throughout the paper I use the terms immigrant and immigrant interchangeably to refer to individuals who were born outside of Spain, regardless of their nationality.

⁹Male and female immigrant population decreased by 500 000 thousand during this period.

¹⁰2014 was the last year where the Spanish economy had a negative growth. I end the main analysis at this year, in order to alleviate concerns about my results being driven by the positive growth observed in the later years. I provide results including the recovery period for robustness.

native workers located in the labor market. The departure of the immigrants reduces the labor supply in the local labor market which improves employment and wages of natives. These benefits are stronger for natives who have higher substitutability with the departing population. For the groups with higher wage rigidities (e.g. older workers, workers with indefinite contracts), the adjustment takes place at the employment margin.

In my analysis, I use the Spanish Social Security panel, which allows me to track natives over time and across firms, sectors, and regions. I combine this data with individual-level municipal registers that cover 100% of the population to precisely count the number of immigrants residing in each province annually. I measure the intensity of departures across Spanish provinces by the annual decline in the immigrant working-age male population during this period, relative to the entire working-age male population in the previous year.

I find a positive and causal effect of the out-migration of the immigrant workers on the wage and employment growth of natives located in the area. During the 2009-2014 period, a 1% increase in the annual out-migration rate of the immigrant population increases the local native wages and employment by about 2% and 2.4%, respectively. These results are robust to pre-trends, use of different measures of outflows, alternative weights and instruments, the inclusion of time-varying controls which capture the changes in the local demand, the structure of the local economic activity, and composition of the workforce. Given the recession context and the overall decrease in the employment during the period, these effects suggest that immigrant out-migration dampened the negative effects due to the demand shock by slowing employment and wage losses in the local labor market.

Regarding differential effects by skill groups, the out-migration leads to an employment increase for both unskilled and skilled natives. Breaking employment responses further out by age group and gender, my results reveal that groups that have the highest substitutability (e.g., natives under 30) or high elasticity (e.g., natives above 50) benefit the most in terms of employment, which is in line with the standard migration model.

My decomposition of the overall native employment response into different margins of adjustments sheds light to the mechanism. First, the departure of the immigrant population increases the native population in the area, through increased inflows from other areas, while the outflows do not decrease significantly.¹¹ I find that for every ten immigrant that leave an area, three natives enter. This finding is in line with recent literature which finds that population inflows respond more than outflows to economic shocks (Monras, 2018) or increases in the labor supply due to immigration (Dustmann et al., 2017a).

Second, I show that outflow of the immigrant population increases the recruitment of natives who were *not* employed previously, while those who were already employed were not less likely to leave employment. This observation indicates that departure of the immigrant population benefits “outsiders”

¹¹The question of whether an immigration-induced labor supply shock leads to relocation of the natives remain controversial. See for instance, Card (2001); Peri and Sparber (2011); Borjas (2003).

(and not “insiders”) at the employment margin. This finding is in line with Dustmann et al. (2017a), who find that an increase in the labor supply due to the arrival of immigrant workers reduces the employment of “outsiders”, while “insiders” are not affected thanks to labor market institutions. Put together, these findings suggest that employment adjustments to positive (negative) shocks benefit (hurt) the outsiders in rigid labor markets. On the other hand, I find that workers who were already in the labor market (“insiders”) are the only ones who benefit from wage increases.

My findings have implications for multiple strands of literature. First, they relate to the literature showing that mobility of immigrant workers can be a mechanism for equalizing differences across labor markets (Bartik, 1991; Blanchard and Katz, 1992). Due to their higher responsiveness to spatial differences in economic opportunities compared to natives, immigrants have much higher mobility response which “greases the wheels” of the labor markets (Borjas, 2001; Cadena, 2010).¹² My paper is specifically related to the recent literature focusing on the immigrant mobility and its cushioning effects for the natives during a demand shock. In a seminal work on the Great Recession in the US, Cadena and Kovak (2016) shows that the mobility of Mexican workers reduced the incidence of local demand shocks on natives. In a similar study, Basso et al. (2018) also finds that the immigrant workers’ mobility in the Euro Area is strongly cyclical and it reduces the variation of overall employment rates over the business cycle. I advance this literature by using individual-level administrative data to provide the first estimates on the direct impact on wages and employment margins controlling for selection, while accounting for other adjustment margins that it may entail, making it possible to provide a more comprehensive picture of the total effects. I further contribute to the literature by presenting a general equilibrium model that predicts how a decrease in the labor supply due to immigrant mobility affects wages and employment of native workers in the local labor market.

My paper is closely related to a second literature which focuses on the labor market impact of out-migration-induced supply shocks. Despite the extensive literature studying the effects of immigration, evidence on the effects of out-migration is scant. Focusing on the emigration of the Mexicans to the US between 1970 and 2000, Borjas and Aydemir (2007) and Mishra (2007) presented the first econometric results on the effect of emigration on the national wages. Similarly, Elsner (2013) and Dustmann et al. (2015) study the effect of the emigration on the wages of the staying natives in Lithuania and Poland, respectively. My paper differs from these for two reasons. First, I examine the effects of the departure of the immigrant population, on the staying natives in the host country. Second, apart from Dustmann et al. (2015), all of the mentioned studies focus on the wage effects following the skill-cell approach of Borjas (2003). I follow a pure spatial approach and study the effects of out-migration by exploiting

¹²Borjas (2001) argues that the new immigrant arrivals are much more likely to be clustered in those states that offer the highest wages for the types of skills that they have to offer. In a recent paper Cadena and Kovak (2016) show that the effect is not due to new arrivals. Even immigrants who have arrived earlier are willing to move, making them instrumental for equilibrating labor market differences even after the years.

spatial variations in the total outflows.¹³

Third, I contribute to the literature studying the natives' residential choices, or the so called native displacement, following immigration. The literature provides estimates for the displacement of natives due to arrival of immigrants (Borjas, 2003; Peri and Sparber, 2011; Clemens and Hunt, 2017). I complement this literature by providing the first evidence on how the departure of the immigrants generates the exact opposite effect found in the immigration literature by attracting natives into the area.

The remainder of the paper is organized as follows. The next section provides a theoretical model that allows understanding the impact of a decrease in labor supply due to outflow of immigrants on local labor markets. Section 3.3 gives some background on the economic crisis in Spain and the response of the immigrant population during this period. Section 3.4 describes the data and presents some descriptive statistics. Section 3.5 presents the empirical strategy and addresses identification issues. Section 3.6 and Section 3.7 provides the main results and robustness tests, respectively. Section 3.8 discusses the timing of the adjustment. Section 3.9 provides results for heterogeneous groups, while Section 3.10 discusses the underlying adjustment mechanisms. Section 3.12 concludes.

3.2 An Equilibrium Model with Heterogeneous Labor Supply, Demand Shock and Wage Rigidities

In order to help the interpretation of my empirical findings, I commence by setting out a theoretical framework. I construct a simple aggregate model of an economy where the workers are differentiated by their place of birth (native and immigrant) as well as their education (skill) levels. I also allow exogenous variations in A which captures the changes in the local demand. This structure allows me to examine the wage and employment effects due to changes in the labor supply driven by the outflow of the immigrant population.

The model aims at providing a simple framework, where the labor is the only production input, and output is an interaction of this input with A , which is both TFP and unit price of the output. It is assumed that physical capital is nationally mobile (its supply is perfectly elastic), that each single region is too small relative to the national labor market and returns to physical capital are equalized across locations.

I start out with a fully competitive labor market as a benchmark, and allow for wage rigidities in a second step. Similar to Dustmann et al. (2017a), I allow the labor supply responses of natives, or the degree of wage rigidity, to vary across skill or demographic groups.

¹³Very recently the immigration literature has turned to the labor market effects of immigration restrictions. Clemens et al. (2018) for instance study the effects of the exclusion of Bracero workers in the US in 1964, on the wages and employment of natives in the agricultural sector. They find that the reduction in the entry of seasonal Mexican workers did not increase the wages and the employment of natives in the industry, as the reduction in the labor supply was compensated by an increased in technology adoption in the sector. Similarly, Lee et al. (2017) study the forced repatriation of Mexicans in the US between 1930-1940, and find small decreases in native employment and increases in native unemployment. These effects are not however statistically significant across all specifications.

3.2.1 Set-up

Production Function

The output Y (homogeneous and perfectly tradable) in a specific area is the product of labor L and total factor productivity A .

$$Y = AL \tag{1}$$

Following the literature, I assume that labor L is a nested CES function of skilled (S) and unskilled (U) labor L_g where $g = U, S$.

$$L = \left[\theta_U L_U^\beta + \theta_S L_S^\beta \right]^{\frac{1}{\beta}} \tag{2}$$

where θ_U and θ_S are the productivity levels of unskilled (less than tertiary education) and skilled workers (tertiary education or above). The elasticity of substitution between the two skill groups equals $\sigma = \frac{1}{1-\beta}$, with $\beta \leq 1$.

This representation implies two types of simplifications. First, I assume that the relevant split in terms of production abilities is between college and non-college-educated workers. This is consistent with the previous literature which finds high substitutability between workers with no schooling and high school degree, but small substitutability between these and workers with college education (Card, 2009; Ottaviano and Peri, 2012). Second, for simplicity I omit further classification into age groups, considered as imperfectly substitutable skills (Borjas (2003); Ottaviano and Peri (2012); Docquier et al. (2014)).

I distinguish between natives and immigrant within each skill-specific labour aggregate, L_g^N and L_g^S . Similar to Dustmann et al. (2017a), I assume that within each skill group g , natives (N) and immigrant (M) are perfect substitutes in production, which gives¹⁴:

$$L_g = L_g^N + L_g^M \tag{3}$$

Labor Demand for Natives

Each region is a single labor market. Assuming that firms are price takers in the labor and product market, firms choose labor such that marginal costs equal the marginal products of each type of worker.

¹⁴In my theoretical model, I assume perfect substitution between natives and immigrants within the same skill group, similar to Docquier et al. (2014); Dustmann et al. (2017a). Although elasticity of substitution is almost perfect within refined cells (see Ottaviano and Peri (2011)) it still remains a strong assumption. For this reason, in the empirical section I estimate wage and employment responses for different skill groups to the overall decrease in labor supply due to immigrants' mobility. This means that in my estimation procedure, I do not allocate immigrant workers to skill groups based on their observed skills (Dustmann et al., 2013). Other models have emphasized the role of complementarity within education groups as well as upgrading and specialization of native workers in response to immigrants and have found null or positive wage effects Card (2009); Peri (2012); Ottaviano and Peri (2012). The parameter I estimate will give the aggregate of both complementarity and substitution effects. Complementarity and substitutability (and also the elasticities) between natives and immigrants depend on the period of analysis. For instance, in the short-term, one would expect the substitutability to be stronger while the complementarity is more likely to play out in the medium/long-term.

I derive the marginal productivity for workers of both skills (w_S and w_U) by substituting Equation 2 into Equation 1 and taking the derivative with respect to the total quantity of labour L_S and L_U . This yields the labour demand for each type of worker:

$$w_g = A(\theta_g)L_g^{\beta-1}L^{1-\beta}$$

In Appendix 3.13.1, I take the logarithm of the demand functions for each type and derive the firm's change in the demand of native workers from skill group g , $d\log L_g^N$, due to overall immigration-induced change in the labor supply ($d\log L_g^M$) and the demand ($d\log A$)¹⁵ and obtain the following:

$$d\log L_g^N = \frac{1}{\gamma}d\log w_g - \frac{(1-\beta)}{\gamma}(S_{g'})[d\log L_{g'}^N\theta_{g'}^N + d\log L_{g'}^M\theta_{g'}^M] - \frac{1}{\gamma}d\log A + [(1-\beta)(S_g-1)]\left(\frac{\theta_g^M}{\gamma}\right)d\log L_g^M \quad (4)$$

where g' denotes the other skill group, $\gamma = \frac{1}{[(1-\beta)(S_{g'}-1)]\theta_g^N}$ is the (negative) slope of the aggregate labor demand curve, S_g denotes the contribution of labor type g to the total labor aggregate and θ_g^N and θ_g^M denote the share of workers of skill group g (in head counts) among natives and immigrant (i.e. $\theta_g^N = \frac{L_g^N}{L_g}$ and $\theta_g^M = \frac{L_g^M}{L_g}$).

Suppose that g indexes unskilled labor and g' skilled labor. Equation 4 demonstrates that in the absence of any wage response to immigration (i.e. $\frac{d\log w_g}{d\log L_g^M} = 0$), unskilled native employment declines by the rate $(1-\beta)(S_g-1)\left(\frac{\theta_g^M}{\gamma}\right)$. The equation also shows that a decline in the wage of unskilled labor in response to immigration (i.e., $\frac{d\log w_g}{d\log L_g^M} < 0$) will dampen the employment response of the unskilled, as the slope of the demand curve γ is negative. An increase in the labor demand for other skill groups, would increase the demand for unskilled native employment. (i.e., $\frac{d\log L_{g'}^N}{d\log L_{g'}^M} > 0$, $\frac{d\log L_{g'}^N}{d\log L_{g'}^M} > 0$). Finally, a positive demand shock will also increase the labor demand for natives (i.e., $\frac{d\log L_g^N}{d\log A} > 0$).

Labor Supply

Labor supply of natives and immigrants constitute the total supply. Following the literature¹⁶, I make the simplification that all working-age immigrants supply a constant amount of labour ($\phi_M > 0$) so that total employment of immigrants is given by $L_g^M = \phi_M M_g(A^\alpha)$, where M_g denotes the total number of working-age immigrant population.¹⁷ The size of the immigrant population is function of A, with an elasticity $0 < \alpha \leq 1$.¹⁸

¹⁵Demand for native workers depends on total labor demand and labor supply of the immigrant workers. I include exogenous variation of A in order to capture the changes in the labor demand for native workers that is due to shifts in total labor demand. Thus, the changes in employment of natives is due to both changes in the total demand and the labor supply of immigrant migration.

¹⁶See for instance, Borjas (2003); Docquier et al. (2014); Dustmann et al. (2017a).

¹⁷The goal of this paper is to analyze the effects of a change in immigrant supply on natives' labor market outcomes. This definition implies that a certain percentage change in immigrant population translates into the same percentage change in immigrant employment. See Docquier et al. (2014) or Dustmann et al. (2017a) for a similar simplification.

¹⁸For simplicity, variations in A impact immigrant population of both education groups equally.

Using N_g to denote the (fixed) number of natives who could potentially supply labor to the local labor market, the local labor supply function for skill group g is:

$$L_g = L_g^N + L_g^M = N_g f_g(w_g, w'_g) + L_g^M \quad (5)$$

Local labor supply of natives depends on skill-specific wages in the market under consideration (w_g) and other local labor markets (w'_g). The local labor market elasticity for natives, which I allow to vary by skill group, is then given by $\eta_g = \frac{w_g}{\partial w_g} \frac{\partial(N_g f_g(w_g, w'_g))}{(N_g f_g(w_g, w'_g))}$. Note that η_g is the local labor market elasticity for natives, which varies by skill group. It captures various potential adjustment mechanisms such as moving into and out of non-employment, internal migration of workers between areas, or entries into and exits from the labor force. These adjustment margins may have different importance for different types of workers and thus help explain why some groups respond more elastically than others.

From the labor supply function 5, it follows that (see Appendix 3.13.2 for details):

$$d \log(L_g^N) = \eta_g d \log w_g \quad (6)$$

3.2.2 Equilibrium Effect of Migration and Demand

Competitive Equilibrium with Flexible Wage

In a competitive equilibrium, quantities supplied must equal quantities demanded. The intersection of the demand and supply curve, determine the skill-specific wages and employment in the local labor market.

The equilibrium wage and employment responses are determined by the two skill-specific labor demand curves:

$$d \log w_S = (\beta - 1) d \log L_S + (1 - \beta) d \log L + d \log A \quad (7)$$

$$d \log w_U = (\beta - 1) d \log L_U + (1 - \beta) d \log L + d \log A \quad (8)$$

and two skill-specific supply curves:

$$d \log L_S^N = \eta_S d \log w_S \quad (9)$$

$$d \log L_U^N = \eta_U d \log w_U \quad (10)$$

By substituting Equation 9 and 10 into Equation 7 and 8, and rearranging them, I derive the equi-

librium employment response as (see Appendix 3.13.3):

$$\begin{aligned}
d\log L_g^{N*} &= -\frac{(1-\beta)S_{g'}\theta_g^M\eta_g}{1+(1-\beta)S_{g'}\eta_g\theta_g^N+(1-\beta)S_g\eta_{g'}\theta_{g'}^N}d\log L_g^M \\
&+ \frac{(1-\beta)S_{g'}\theta_{g'}^M\eta_{g'}\eta_g}{1+(1-\beta)S_{g'}\eta_g\theta_g^N+(1-\beta)S_g\eta_{g'}\theta_{g'}^N}d\log L_{g'}^M \\
&+ \frac{(1+(1-\beta)S_g\eta_{g'}\theta_{g'}^N)\eta_g}{1+(1-\beta)S_{g'}\eta_g\theta_g^N+(1-\beta)S_g\eta_{g'}\theta_{g'}^N}d\log A
\end{aligned} \tag{11}$$

Since $\beta \leq 1$, $S_{g'} \geq 0$, $S_g \geq 0$, the denominator, and the numerators are always positive. Thus, an increase in the number of immigrants of the skill group g would have a negative impact on the employment of natives of the same skill group (g) due to substitution. An increase in the number of immigrants in the other skill group (g') would be positive due to complementarities between the two groups.

A negative demand shock (captured in A) would impact the equilibrium through multiple channels:

$$\frac{d\log(L_g^{N*})}{d\log(A)} = \underbrace{\frac{\partial \log L_g^N}{\partial \log A}}_{(-)} + \underbrace{\frac{\partial \log L_g^N}{\partial \log L_g^M}}_{(+)} \times \underbrace{\frac{d\log L_g^M}{d\log A}}_{(-)} + \underbrace{\frac{\partial \log L_g^N}{\partial \log L_{g'}^M}}_{(-)} \times \underbrace{\frac{d\log L_{g'}^M}{d\log A}}_{(-)}$$

A decrease in A will:

- decrease $d\log L_g^N$ (i.e., $\frac{\partial \log L_g^N}{\partial \log A} < 0$)
- decrease the labor supply of immigrants for the skill group g (i.e., $\frac{d\log L_g^M}{d\log A} < 0$), which increases the demand for native labor of the same skill group (i.e., $\frac{\partial \log L_g^N}{\partial \log L_g^M} > 0$).
- decrease the labor supply of immigrants for the skill group g' (i.e., $\frac{\partial \log L_g^N}{\partial \log A} < 0$), which decreases the demand for native labor of the skill group g' (i.e., $\frac{\partial \log L_g^N}{\partial \log L_{g'}^M} < 0$)

The final effect will be the sum of all these forces. A decrease in A will decrease the labor demand for natives, yet it will be dampened by the positive effect due to decreased competition from the immigrant population of the same skill group. In reverse, in case of a positive demand shock (i.e. A is positive), an increase in the labor demand for natives will be slowed by the increase in the competition due to the increased presence of immigrants from the same skill group.

Wage Rigidities

The equilibrium above assumes the flexibility of wages. However, in the context of Spain, wages are rigid and thus have low cyclicality especially in the short-run (see., e.g., Bentolila et al., 2012b; De la Roca,

2014). Moreover, the degree of rigidity may vary across sectors, skills, tenure, type of job contract or occupation (Card and Kramarz, 1996; De la Roca, 2014; Font et al., 2015). These labor market rigidities, while protecting some native workers from negative demand shock or immigrant competition, can also increase the employment response (Angrist and Kugler, 2003; Dustmann et al., 2017a).

If the decline of wages is constrained due to labor market rigidities, then the wages cannot fall by as much as the equilibrium wage response given by the labor supply (Equation 5) or the equilibrium (Equation 11). This would create a demand-side constraint in the market, and in consequence generate an oversupply of (native) workers who would like to work for the current wage rate, but cannot find a job. In this case, the wages would be determined exogenously depending on the wage rigidity for the group, and the employment response of natives would be determined by the labor demand function (Equation 4).

The differences in labor supply responses and the degree of wage rigidities can generate "perverse" effects where the group experiencing the largest shock may not be the one suffering the largest changes in wages or employment.¹⁹

3.3 Spanish Context

This section provides background on the Spanish context by summarizing the evolution of the Great Recession, providing some stylized facts about the immigrant population before the crisis and how this population adjusted to the crisis. In the final part of this section, I discuss differences in the mobility patterns between immigrants and natives.

3.3.1 The Great Recession

Spain was hit by two shocks: the end of the speculative bubble of the construction sector in Autumn 2007 and the global financial shock in September 2008.²⁰ The negative shock in the construction sector reversed the positive trend in the employment observed until the crisis. The global financial shock triggered a rapid increase in the unemployment rate.

Both shocks impacted the Spanish labour market very severely.²¹ The deterioration in the labour market following the crisis in 2008 unfolded in three stages. In the first stage of the crisis, very sharp job losses took place with annual decreases that exceeded 6% of the total employment, driven by the sharp decline in GDP. Over the course of 2010, the decline rate in employment tended to soften, helped by some recovery in GDP. However, this recovery was not sustained and led to a double-dip recession

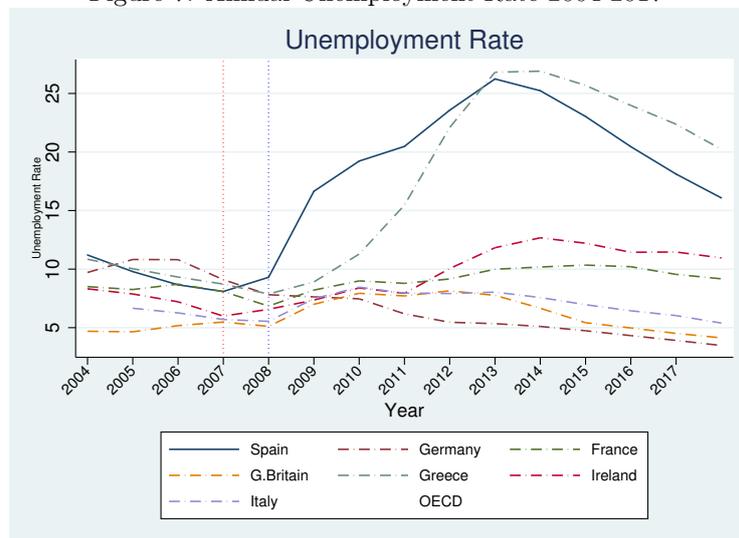
¹⁹See Edo (2016); Verdugo (2016); Dustmann et al. (2017a) for a similar results in other rigid European markets.

²⁰Prior to the crisis, Spain has experienced one of the most important housing booms among developed economies. See Gonzalez and Ortega (2013); Akin et al. (2014); Sanchis-Guarner (2017) on the topic.

²¹According to Akin et al. (2014), Spain was hit more severely than other developed countries due to the joint presence of excessive dependence of the real estate industry and soft credit standards which were applied during the boom years. Moreover, Moral-benito (2018) argues that too much credit was given to firms with high real-estate collateral, especially in municipalities with higher housing prices growth. This rendered Spanish firms much more reliant on bank credit than other similar countries and vulnerable to credit supply shocks (Bentolila et al. (2017)).

from early 2011 that once more intensified job losses. In this relapse employment accumulated a further decline of around another 8% taking it to its cyclical trough in late 2013 to 26%, which was slightly lower than Greece (Figure 7). Since 2014, the economy started recovering by recording positive employment growth.

Figure 7: Annual Unemployment Rate 2004-2017



Data Source: OECD

Job losses were highest among workers with temporary contracts (more than 1.7 million between 2008 and 2013), due to high reliance on the use of such contracts in the Spanish economy.²² Although the losses were concentrated among temporary workers in the first stage of the crisis, very soon the dismissals of workers with open-ended contracts reached historical levels (Malo, 2015).

Spain is a country characterized by a rigid labour market with very low wage cyclicality (De la Roca, 2014; Font et al., 2015). This rigidity, due to various institutional features such as strong wage indexation to inflation, dual market structure or the collective bargaining system, makes it very difficult for wages to adjust to the economic cycles (Bentolila et al., 1994; Messina et al., 2010; De la Roca, 2014; Font et al., 2015). More importantly, the rigidity is found to be even stronger during economic downturns limiting wage adjustments in recessions (Font et al., 2015).

As the contraction in the economic activity deepened, the labour costs continued to increase. The low association with the cyclical conditions in the labor market and real wages, limited the use of wage reductions by the firms as a channel to adjust to the lower demand in the market. When the crisis initiated in 2008, the wages failed to react to the strong deterioration of the labor market, and saw a

²²Some seasonal activities (such as tourism) and/or per-task activities (such as construction) have a higher share of temporary contracts compared to their total employment. However, they are not the main reason for the high reliance on temporary contracts. In fact, temporary contracts are more widespread in Spain than in other countries irrespective of the sector, industry or occupation (Malo (2015)). Also see Bentolila et al. (2012a) for a discussion on the role of temporary contracts in the increase in the unemployment rates in Spain.

rise in real wages was observed in 2008-2009, even after controlling for the strong composition effects in employment (Puente and Galán, 2014). Although some reduction in real wages was observed in 2012 and 2013, negative inflation reduced the scope for further real wage adjustment.

The rigidity of wages in the initial phases and the small adjustment in the following years forced the firms to adjust to the increase in the labor costs and the decrease in demand via employment margin, which caused the unemployment rates to triple, generating a high employment-GDP elasticity (Bentolila et al., 2012a).²³²⁴

3.3.2 Immigrant Population in Spain

Before the Crisis

The decade between 1998 and 2008 has been characterized by one of the most significant immigration episodes in recent history among the OECD countries.²⁵ Until 2009, Spain received an average of almost half a million immigrants annually, thus becoming the second-largest recipient of immigrants in absolute terms in the OECD after the United States (Arango, 2013). As can be seen in Figure 8, the immigrant share in the total population increased from 1.6% in 1998 to 12.1% in 2009, reaching to 5.6 million.²⁶ Overall immigrants to Spain had higher average years of schooling than natives, and thus their arrival contributed to increasing the average level of human capital in the country.²⁷ Prior to the crisis, the average schooling and age of entering immigrants decreased (Fernández-Huertas Moraga, 2014). During this period, the employment share of immigrants in construction, services and domestic help rose markedly (Gonzalez and Ortega, 2011; Farré et al., 2011; Gonzalez et al., 2009).²⁸

²³The adjustment through employment margin is not unique to the Spanish context. Cadena and Kovak (2016) and Rothstein (2012) argue that during in the US, changes in average wages were relatively small compared to the substantial changes in employment during the Great Recession.

²⁴Wage Dynamics Network (WDN) survey, conducted by the European Central Bank across the Eurozone, offers important insights into the effects of the crisis on firms and their adjustment during the 2010-2013 period. Survey results show that due to the wage rigidities, firms of all sizes and sectors made adjustment at the employment margin rather than wage margin. See Section 3.13.4 for more details.

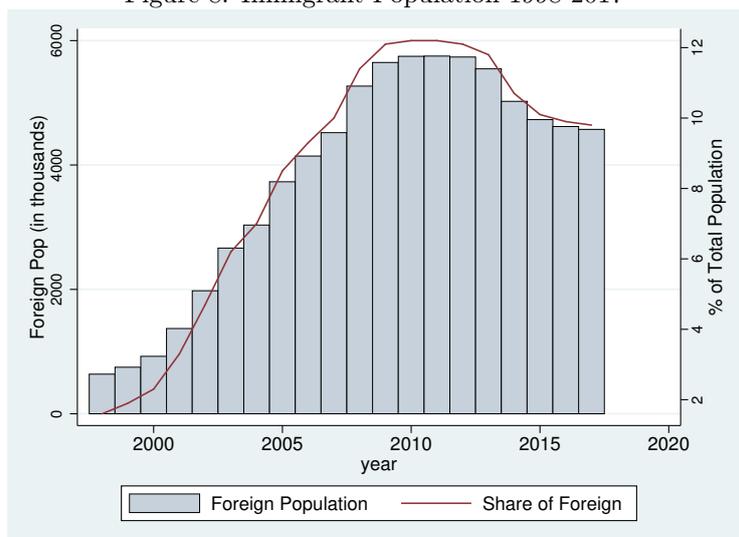
²⁵The arrival of Russian Jews to Israel following the disintegration of the Soviet Union in the early 1990s is the other important immigration episode observed in recent history. See Friedberg (2001) on the topic.

²⁶This figure is calculated based on the nationality of the individual. Naturalisation rates are high in Spain, especially for those originating from Latin American countries. If, as opposed to nationality, one calculates this figure based on the country of birth of the individual the number is much higher. For instance in 2009, there were 6.4 million immigrant people, which makes the share of immigrant in the population 13.7%. Throughout the paper, I use immigrant and immigrant interchangeably, to indicate individuals born outside of Spain.

²⁷Fernández-Huertas Moraga (2014) shows that both before and after the crisis, the average years of schooling of immigrants always remained above those of natives.

²⁸Appendix Figure 19 plots the evolution of immigrant working-age male population, as a share of total working-age male population by skill group. Between 2005 and 2009 period, the share of immigrants in the total low-skilled population increased from 12.3 in 2005 to 17.9 in 2009. During this time, share of immigrants in high-skilled population increased from 10.9 to 12.2.

Figure 8: Immigrant Population 1998-2017



The figure shows the evolution of immigrant population (both male and female, all age groups), and its share as of total population. Data: Authors own calculation. INE

Spain has an extremely diverse immigrant population (de la Rica et al., 2014). While Romania is currently the main country of origin, followed closely by Morocco, the largest region of origin is South and Central America and, particularly, Ecuador and Colombia, which accounted for over one quarter of Spain’s immigrants.

A substantial portion of immigration to Spain was driven by labor market motives due to the strong economic growth (de la Rica et al., 2014).²⁹ In addition to the economic pull factors, cultural and linguistic factors also played a role in shaping Spain’s immigration experience.³⁰ In addition to the cultural proximity, the special arrangements that allowed citizens of the former colonies to enter Spain without a visa generated large migration from Latin America (Bertoli and Fernández-Huertas Moraga, 2013, 2015).³¹

Spain also receives a substantial amount of family-based and retirement migration. Retirement migration is mainly composed of immigrants from the United Kingdom, Germany, France (which together account for two-thirds of foreigners from the EU-15 in Spain), and other northern European countries who are attracted to Spain by the country’s temperate climate, among other factors.

Before the crisis in 2007, the average immigrant share in Spanish provinces was 8 percent. As can be seen in the map in Appendix 23, although immigrants settled across Spain, most of them settled in large cities or coastal provinces. For instance in 2007, the immigrant share in the total population

²⁹Between 1995 and 2007, the Spanish economy experienced the longest expansion in its recent history, where the real GDP grew above 3.5% per year (Moral-benito (2018)).

³⁰In a cross-country study, Adserà and Pytliková (2015) underline the importance of linguistic similarity in shaping international migration flows.

³¹This rapid increase in the immigrant population generated academic interest in the labor market impact of immigration. See for instance Amuedo-Dorantes and De La Rica (2005); Carrasco et al. (2008); Gonzalez and Ortega (2011); Amuedo-Dorantes and de la Rica (2011) for labor market impact of immigration. See also de la Rica et al. (2014) for a more comprehensive review of the literature on the immigration wave experienced by Spain over the last decade on other dimensions (i.e. housing, immigrant assimilation, inflation rate etc.).

of Barcelona and Madrid were 12.5 and 14.5 percent respectively. In addition to clustering in large cities, many immigrants settled in coastal provinces with high levels of tourism and European retirees. Provinces such as Tarragona, Castellon, Almeria, Girona, the Balearic Islands, Alicante had immigrant shares above 15 percent. On the other hand, there were many provinces with extremely low levels of immigration: more peripheral provinces such as Coruña, Asturias or Lugo in the north; Cordoba, Jaén, Sevilla or Cádiz in the south; and provinces in central Spain all had immigrant shares that were 4-5 percentage points below the national average.

Immigrants to Spain integrated quickly to the labor market and exhibited a higher labor force participation rate compared to their native counterparts. This sets Spanish experience apart from what has been observed in the other European host countries (De La Rica et al., 2015). Especially immigrants from Latin American countries integrated faster: in the first year of arrival, their employment rates were equal to that of comparable natives (Amuedo-Dorantes and De La Rica, 2005).³²

Focusing on the occupational distribution of immigrants, Amuedo-Dorantes and De La Rica (2005) found no occupational segregation between EU immigrants and natives³³ while finding evidence for occupational segregation of non-EU immigrants into low-skill occupations. Over time, Eastern Europeans and Latin American immigrants experienced an improvement in their labor market conditions by moving up to better paid occupations, while no such progress was observed for Africans.³⁴ Female immigrants on the otherhand were confined into a few “niche jobs” such as domestic service or childcare.

The Crisis: 2007-2014

During the crisis, immigrants, especially male immigrants, suffered higher unemployment rates than native workers (See Appendix Figure 15).³⁵ There are many reasons why immigrants in Spain were hit harder. First, immigrants were concentrated in sectors which are more sensitive to the business cycle.³⁶ Secondly, immigrants were more likely to have less secure contractual agreements. Immigrants were more likely to hold temporary contracts prior to the crisis which makes them more vulnerable to firing (Fernandez and Ortega, 2008).³⁷

The crisis caused a decrease in the immigrant population, driven by changes in immigrant inflows

³²The authors also find that Spanish-speaking immigrants specialized in occupations that are intensive in communication tasks, similar to the natives.

³³Note that this study was published before the entry of Romania and Bulgaria to the EU. Thus only includes EU nationals who were member prior to 2007.

³⁴Izquierdo et al. (2009) presents a similar picture regarding the wage gap between immigrants and natives. They find that earnings assimilation is much faster for South-American and EU immigrants compared to Africans.

³⁵There is ample evidence that immigrants’ employment and earnings are more sensitive to business cycle fluctuations than natives’. See for instance Orrenius and Zavodny (2010) and Dustmann et al. (2010) for evidence from the US and Europe, respectively.

³⁶They were recruited in sectors that played a key role during the expansion period, such as construction, wholesale and hotels and restaurants where up to 50% of the immigrants were employed.

³⁷The fact that the changes in the business cycles affect immigrants differently has been explored in other country contexts. Dustmann et al. (2010) find larger unemployment responses to economic shocks for immigrants relative to natives within skill groups in the UK and Germany, especially for non-OECD immigrants. Focusing on the Great Recession in the US, Orrenius and Zavodny (2010) also conclude that immigrants, especially the low-skilled, have higher sensitivity to the business cycle than natives;

and outflows.³⁸ Spanish Statistical Institute (INE) provides Residential Variation Statistics (EVR, or *Estadística de Variaciones Residenciales*, in Spanish), a micro-data which records all individual moves originating or ending in Spain based on the Municipal Register of Population (*Padrón Municipal de Habitantes*, in Spanish). In Figure 9, I use this data to plot the annual volumes of international arrivals and departures for immigrant working-age male population. As can be seen from the figure, starting from 2007, entries dropped dramatically from more than 400 thousand per year to below 200 thousand in 2009, going all the way down to 40 thousand in 2013.³⁹ This significant and immediate drop in labor related entries was due to two reasons. First, unlike most of the other European countries where drop was less pronounced and gradual, migration flows in Spain have always depended highly to economic cycles (OECD, 2009). Due to the contraction in the economy, Spain's appeal as a destination decreased.⁴⁰ Second, Spanish government took action to reduce the labor related entries (see Appendix 3.13.8 for more details). In parallel to decrease in inflows, departures also started increasing due to contraction in the economic activity.⁴¹ In addition to voluntary departures, Spanish government also took action to encourage departure of immigrants (see Appendix 3.13.8 for more details).⁴²

³⁸It is important to understand whether, for example, a net decrease in the stocks, is due to increased outflows, decreased inflows or the sum of both factors. It is important to understand from which margin the effect comes from as it may have differential effect on the labor market.

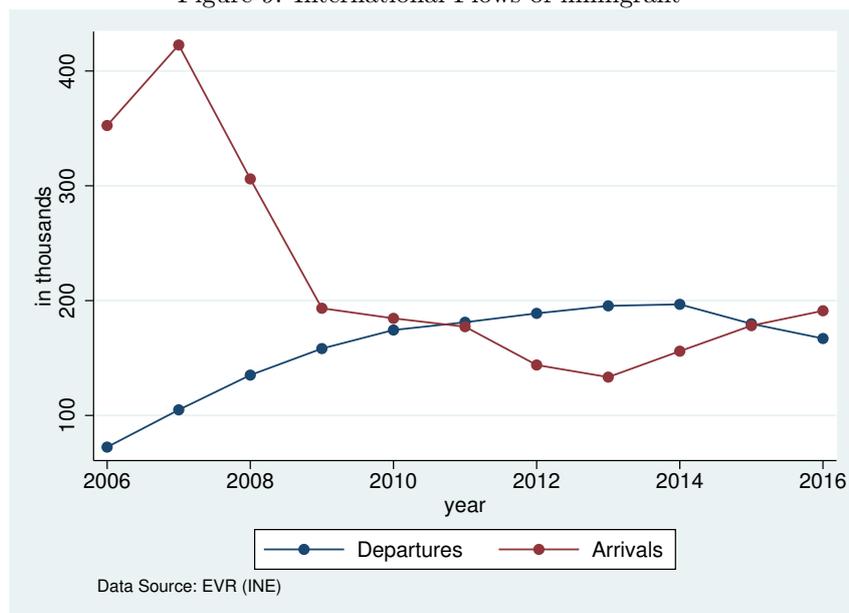
³⁹Migration outflows started to increase in 2007 when GDP growth in Spain started to decelerate. Still, in the first years into the crisis, although labor motivated entries decreased, the number of immigrants continued to grow, mainly due to family reunification and existing migration networks.

⁴⁰For instance, the number of foreigners residing abroad who are offered jobs in Spain — a process known as *contratación en origen* declined from 45 995 in 2006 to 4 429 in 2009.

⁴¹This data has been used in the literature to consider both international and domestic moves. See for instance Bertoli and Fernández-Huertas Moraga (2013); Gil-Alonso et al. (2015); Melguizo and Royuela (2017); Amuedo-Dorantes et al. (2018)

⁴²A reasonable question is whether the economic crisis impacted flows differently for those who are from EU countries, thus benefiting from free mobility, vs. those who are not. I discuss these issues in Appendix Section 3.13.7 further and show that there are no striking differences between the mobility patterns of the two groups.

Figure 9: International Flows of immigrant



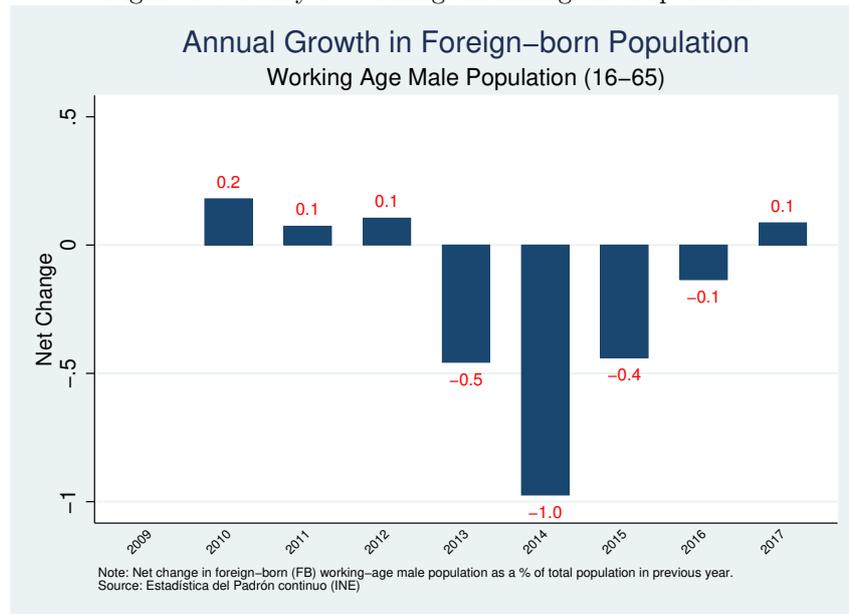
Note: The figure presents the total inflows and outflows for the immigrant working-age male between 2006-2016. Data source: EVR

Between 2009-2014, working-age male immigrant population saw a net decrease of 280 thousand.⁴³ Figure 10 shows the annual net outflow of the working-age male immigrant population as a share of the total working age population (Spanish and immigrant) the year before. It can be seen that the departure of the immigrant created a labor supply shock between 0.1-1.4% annually.⁴⁴ Between this period, the total net outflow of working-age immigrant male population caused a reduction of 2.1% in the total male population across Spain.

⁴³This is consistent with the findings of Fernández-Huertas Moraga (2014), who show that the female share in among immigrants started increasing during the crisis.

⁴⁴The net decrease was even stronger as share of the immigrant population. During this period 2.8% of the immigrant population left annually. The net decrease between 2009 and 2014 corresponded to 10.1% reduction in the immigrant male population in 2009. Native male population of working-age only decreased by 3.2% during this period.

Figure 10: Yearly Net Change in Immigrant Population



Both in terms of numbers and as a share of the group’s total population in Spain, most emigrants were Europeans (of new member countries) and South Americans. Although Africans also emigrated their share in the total outflows was much smaller. In terms of skills, most of the outflow during the period happened for those who were low-skilled due to negative selection of immigrants who left after 2008 (Fernández-Huertas Moraga, 2014; Izquierdo et al., 2016, see also Appendix Figure 19). Similarly, departure of the younger immigrants also increased the average age of immigrant population in Spain.

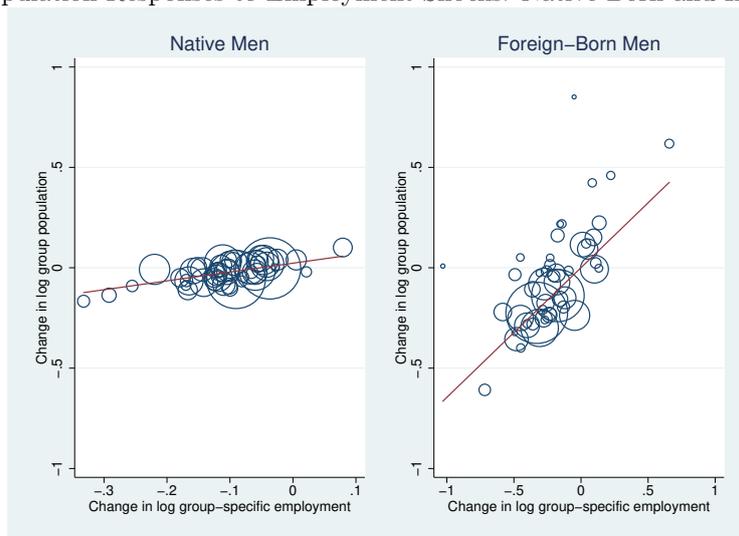
3.3.3 Immigrant vs. Native Mobility

Immigrants are more mobile than native-born workers across regions, industries and occupations (Borjas, 2001; Orrenius and Zavodny, 2007). Recent literature shows that native-born population are less sensitive to labor demand shocks and respond much less by geographic mobility compared to immigrant counterparts (Schündeln, 2014; Cadena and Kovak, 2016; Bartik, 2017; Basso et al., 2018). These differences can be due their observable demographic characteristics (i.e. age, education, family structure, home ownership), but also unobservable characteristics (i.e. self-selected group of people with high levels of labor force attachment and a greater willingness to move long distances to encounter more favorable labor market conditions). The difference in the responsiveness between natives and the immigrant is especially high within the lower-skilled (Schündeln, 2014; Cadena and Kovak, 2016; Bartik, 2017; Basso et al., 2018).

I start by establishing whether the difference in mobility between natives and immigrants exists in Spain. Figure 11 shows scatter plots for working-age native-born and immigrant men and compare their

mobility. Each circle represents a province where the size is proportional to the province population. The x-axis shows the change in log employment, and the y-axis shows the change in log population for the relevant group. Note that the changes in the log population can be due to both internal and international mobility. The figure demonstrates that immigrant workers respond much more strongly to local labor demand shocks than natives. The immigrant population in hard-hit areas move out, while natives remain in the area.

Figure 11: Population Responses to Employment Shocks: Native-Born and Immigrant Men



The figures plot the changes calculated as the long difference in logs from 2009 to 2014 for each province. The x-axis is the group-specific change in employment and y-axis is the change in population. Observations are weighted by the group population in 2009.

Data: Labor Force Survey (EPA), INE

These figures show that native population was less responsive to shocks compared to immigrant population. The higher mobility of the immigrant compared to the natives in Spain has also been confirmed by earlier work focusing on the internal migration in Spain (David and Javier, 2009; Hierro and Maza, 2010; Gil-Alonso et al., 2015; Gutiérrez-Portilla et al., 2018b; Maza et al., 2019). More recent work focusing on individual internal moves in a gravity-type setting, show that immigrant moves have much higher elasticity to labor market conditions (Gutiérrez-Portilla et al., 2018a; Melguizo and Royuela, 2017; Maza et al., 2019).⁴⁵ Maza et al. (2019) further show that while in the pre-crisis period both natives' and immigrants' moves were also partially motivated by amenities (e.g. temperate climate, sunny days), these amenities lost their importance for foreigners during the crisis.

In Appendix 3.13.6, I further explore the issue by decomposing the mobility by the skill group to see if the elasticities differ depending on the skill group. Similar to findings of Cadena and Kovak (2016) for

⁴⁵The low sensitivity of native flows on changes in unemployment is not a new phenomenon in Spain. For similar analysis for earlier periods see Antolin and Bover (1997); Bentolila and Dolado (1991).

the US, I find that the native population in Spain is much less sensitive to the demand shocks. While high-skilled natives have higher elasticities than low-skilled natives, these elasticities still remain lower than those estimated for low-skilled immigrants.⁴⁶

Despite the low share of native outflows compared to the population, it is important to account for them empirically as they could matter for the re-adjustment of the local labor markets. In Section 3.7.2, I present results controlling for native mobility.

3.4 Data and Summary Statistics

This section presents the data used to estimate the effects of net outflow of immigrants from Spanish provinces on natives' outcomes. After describing the data and selected sample, I provide some descriptive statistics about the immigrants and natives.

3.4.1 Data

Social Security Data

The main data come from Spain's Continuous Sample of Employment Histories (MCVL or *Muestra Continua de Vidas Laborales* in Spanish).⁴⁷ This is an administrative data set with longitudinal information obtained by matching social security, income tax, and census records for a 4% non-stratified random sample of the population who in a given year have any relationship with Spain's Social Security. Individuals can either be working as employees or be self-employed, receiving unemployment benefits or pension.⁴⁸

An individual enters the sample if he registers one day of activity with social security, between 2004-2016 and is kept in subsequent editions. Once in the sample, MCVL records any changes in individual's labor market status or job characteristics (including changes in occupation or contractual conditions within the same firm) since the date of first employment. I combine multiple editions of the MCVL, to construct a panel that has the complete labour market history for a random sample of approximately 4% of all individuals who have worked, received benefits or a pension in Spain at any point since 2004.⁴⁹ By combining multiple waves, enlarge the sample by including individuals who have an affiliation with the Social Security in one year but not in another. This allows me to maintain the representativeness of the sample throughout the study period. Individuals who stop working remain in the sample while

⁴⁶Despite large differences in unemployment rates and economic conditions, labor mobility has been low in Spain especially compared to the European countries (Mulhern and Watson (2009); Bell et al. (2015)). There are many reasons why natives react less to demand shocks. In the case of Spain, the low mobility is partially explained by the safety net provided by the families (Bentolila and Ichino (2008)), or by the presence of welfare state which decrease the incentive for such moves (Bover and Velilla (2005)Amuedo-Dorantes et al. (2018)).

⁴⁷This dataset is distributed by Directorate General of Planning for the Social Security (*Dirección General de Ordenación de la Seguridad Social*) under the Ministry of Labor, Migrations and Social Security (*El Ministerio de Trabajo, Migraciones y Seguridad Social*).

⁴⁸This dataset has been widely used in research on labor markets (e.g. Gonzalez and Ortega, 2011; Bonhomme and Hospido, 2017; De la Roca, 2017; De la Roca and Puga, 2017).

⁴⁹I am able to link individual records across waves as MCVL keeps a unique identifier for individuals across waves.

they receive unemployment benefits or a retirement pension, and drop from the sample when their unemployment benefits run out, die or leave the country permanently.

On each date, I know the individual's labor market status, daily wage⁵⁰, the occupation and type of contract, the establishment's sector of activity at the NACE three-digit level, and the establishment's location if it is located in a municipality with population greater than 40,000 inhabitants. I also obtain individual characteristics such as age, gender, country of birth, nationality, and educational attainment which come from *Padrón* or Municipal Register.⁵¹ Furthermore, by exploiting the panel dimension, I construct precise measures of tenure and experience, calculated as the actual number of days the individual has been employed, respectively, in the same establishment and overall.

This rich administrative data set is well suited for my analysis for multiple reasons. First, MCVL allows me to track individuals across time and space based on their workplace location. By using the individual data, I can construct constant cohorts of workers within demographic and skill-groups, avoiding compositional biases that confound cross-sectional analyses. Second, the large sample size allows me to obtain precise estimates of out-migration on wages and employment even for specific subgroups. Third, the longitudinality of the data allows me to investigate whether an increased outflow of workers drives the employment effects into other areas or non- or unemployment, or by a decreased inflow of workers into the local labor market. Fourth, in addition to information on education, age, tenure, and other individual characteristics, the data include both the citizenship and country of birth, which allows identification of all immigrant workers who have naturalized.⁵²

Sample Restrictions

After combining the social security and income tax records, my monthly panel covers job spells in 2005-2016 for individuals aged 18 and over, born since 1962, and employed at any point between January 2005 and December 2016. This initial sample has 777,593 workers and 75,945,441 monthly observations.

First, I exclude spells for workers who are self-employed because labour earnings are not available during such periods, but still include job spells as employees for the same individuals. I also exclude

⁵⁰The MCVL contains earnings data coming from the social security records and income tax records. The earnings data coming from the social security records that go back to 1980. However, the earnings data are either top or bottom coded for about 13% of observations. Income tax records on the other hand provide uncensored gross earnings starting from 2004. Each source of labour income is matched between income tax records and social security records based on both employee and employer (anonymized) identifiers. The Basque Country and Navarre autonomous regions collect income taxes independently from Spain's national government. As job spells originating from these regions lack uncensored earnings information, I use earnings coming from the social security records after applying necessary wage simulations following de la Roca (2015). I further check the robustness of my results by using uncensored earnings information coming from tax records.

⁵¹Since 2009 the Ministry of Education directly reports individuals' highest educational attainment to the National Statistical Institute and this information is used to update the corresponding records in the Continuous Census of Population.

⁵²This feature is valuable for three reasons: first, Spain permits dual citizenship under limited circumstances. Although the requirement to have a single nationality is waived for natural citizens of many countries, nationals of some countries are still required to renounce their original citizenship when they become Spanish. Second, Spain this is has high naturalization rates. As shown in Ródenas et al. 2017, in 2014, 24.7% of foreigners residing in Spain were Spanish citizens. Those with dual citizenship can be registered either as a immigrantnational as well as Spanish both in the municipal as well as social security records. Third, as pointed in Ródenas et al. 2017, naturalized immigrant have higher probability of emigration back to their home country during the crisis. These factors make the naturalisation an important measurement issue given the research question in this paper.

job spells where the workplace location is not available for municipalities with population below 40,000 (which concern rural areas and three small urban areas). I also exclude Ceuta and Melilla given their special enclave status in continental Africa.

Job spells in agriculture, fishing, mining, and other extractive industries are excluded because these activities are typically rural and are covered by special social security regimes where workers self-report earnings and the number of working days recorded is not reliable (De la Roca and Puga, 2017). Job spells in the public sector, international organizations, and in education and health services are also left out because earnings in these sectors are regulated by the national and regional governments. Apprenticeship contracts and certain rare contract types are also excluded. I drop workers who have not worked at least 30 days in any year and those with missing education data.

I exclude women and immigrants from the sample. I exclude women for three reasons. First, I drop women to provide estimates that can be comparable with other studies on the labor market impact of immigration (De la Roca, 2014; Ortega and Verdugo, 2016; Dustmann et al., 2017a; Edo, 2017; De la Roca and Puga, 2017). Second, despite the important increase in the female employment, their participation rate is still low compared to similar countries. For instance, the female labor force participation rate was 49.1% in Spain, compared with 64.5% in the EU-27 and 72% in the US. The gender gap in employment is also among the highest in the industrialized countries (21% in Spain, 18% in the EU-27 and 10% in the US) (Farré et al., 2011). Finally, I prefer excluding women as their employment decisions depends not only on the labor market opportunities but also on the cost of child-care services and elderly care (Farré et al., 2011; Cortés and Pan, 2018), and on their husbands' employment especially during an economic crisis (Baslevent and Onaran, 2003). In this sense, substitutability and complementarity relationships between immigrants and native workers might be different for men and women (Carrasco et al., 2008).⁵³ Although women workers are excluded in the main results, I present some result in Section 3.9. I leave-out immigrant workers as I am interested in the outcomes of natives.

Finally, I restrict the sample to works ages 25 to 54 between 2009-2014 to ensure that individuals have completed their education and avoid complications related to retirement decisions.⁵⁴ I also only keep those who work full-time. These restrictions reduce the sample to 193,247 native-born workers with 814,197 yearly observations. This means that on average each native is observed for 4 years. I use this sample to calculate province-year averages and changes.

⁵³For example, child-care services provided by immigrant women might have allowed some native women to participate in the labor force as shown in Farré et al. (2011) or Cortés and Pan (2018).

⁵⁴I apply these restrictions to avoid problems with potentially endogenous labor market participation in educational decision for young people and early-retirement decisions for men. See Hunt (2017) and De la Roca and Puga (2017); Ortega and Verdugo (2016) for a similar approach.

Mobility Data

Outflows from provinces are measured using microdata from the Municipal Register of Population (*Padrón Municipal de Habitantes*), which is the official population registry of municipalities. According to law (*Ley de Bases de Régimen*), anyone living in a Spanish municipality is obliged to register upon arrival in the country, and to de-register upon departure. Arriving individuals have strong incentives to register as it allows them to enjoy municipal services (such as getting a national ID, drivers permit, passport, proof of residence) and grants them access to regionally-provided ones (such as education and health).⁵⁵ For instance the residents are assigned to schools and hospitals based on the residential proximity measured by their official registration. If not registered, individuals do not have access to any of these services. All newborn children are immediately registered before discharge, and deceased persons are removed upon death. The data covers close to 100% of the population (Fernández-Huertas Moraga, 2017).

Immigrants have additional incentives to register, which makes this data particularly important for recording foreigners that are residing in Spain both legally and illegally. Since 2000 (*Ley Organica 4/2000*), regardless of their status, registered immigrants have been entitled to make use of the public health system and education with no risk of detention by the authorities. This incentivize the illegal immigrants to report their presence (Bertoli et al., 2013). Moreover, registration has been used to prove residency in the periodical regularizations (*Ley Organica 4/2000*). Hundreds of thousands of immigrants took advantage of being duly registered in the 2000, 2001 and 2005 amnesties (see Monras et al. (2019) for the impact of the last regularization wave).

Due to this structure, municipal registers provide precise numbers on the immigration and internal moves. However, the numbers are less precise in recording emigration due to few reasons. First, individuals register and de-register on the basis of their planned length of stay in the country (for entries) or the planned length of absence from the country (for exits), so some individuals may leave the country without de-registering if they plan to return shortly. Moreover, some individuals may prefer not to de-register to keep their entitlements associated with residency. Finally, individuals may simply do not think about de-registering as it does not provide any additional benefits.

Since January 2006, the INE corrects this by requiring local authorities to de-register immigrants if they do not confirm their residence within two years.⁵⁶ Once a registration is deleted, it is counted in the official data as a departure to an unspecified destination country. Data from 2006 thus includes not only return and non-return out-migrations for which the departure is registered, but also those resulting from non-renewal of residency within two years.⁵⁷

⁵⁵For more details, see Rodenas Calatayud and Marti Sempere (2009).

⁵⁶This however is only done to immigrant without an EU nationality or a permanent residence permit.

⁵⁷In case of non-renewal of the residence, the deletions are carried out automatically, exactly two years after the day of registration. This means that Padrón records can suffer from measurement error in terms of exact timing of the departure, and depending on the nationality of the departed individual. I discuss further these issues in Appendix 3.13.10 and show

I have access to individual micro-data from the official registry, which provides information on individual's municipality of residence, nationality, the place of birth (municipality if born in Spain, country of birth otherwise), age and sex. Having these details allow me to calculate precise stocks of the immigrant and native stocks flexibly according to my sampling criteria (i.e., age, sex, country of birth). I use the changes in these stocks to account for variations in the labor supply similarly to the papers that study the impact of changes in labor supply (Cadena and Kovak, 2016; Gonzalez and Ortega, 2013; Ortega and Verdugo, 2016; Lee et al., 2017; Sanchis-Guarner, 2017).

I use population stocks to account for mobility of the immigrant population for two reasons. First, it is considered to be a good measure for the number of immigrants living in the country, especially for both those residing legally and illegally, working in formal or informal sector. Second, as I am interested in changes in supply as a share of the total population at the previous period, it is more precise.

The only limit of this data is that it does not allow to distinguish whether the changes in population stocks are due to international or internal mobility. To verify that changes in the population stocks are not driven by either internal or external moves, I use the EVR micro-data and find that internal moves constituted 55% of the moves while external ones 45%.⁵⁸

Using the municipal registers, I calculate the total stocks (both immigrant and Spanish⁵⁹) and the net change in immigrant population between two periods using male population that is of working-age (16-65). I apply these criteria as these stocks provide more precise measures for those involved in the labor market. As a robustness, I also provide results using female population and all the age groups, which provide similar results with smaller magnitudes as expected. Figure 12 shows the total outflows during the period 2009-2014, as a share of the total population in 2009. It can be seen that the net departures of the immigrant corresponded on average to 3% of the working-age population of the province during the period.⁶⁰ It is also important to note that the net outflow of the immigrants was observed in all the provinces (apart from Guipuscoa).

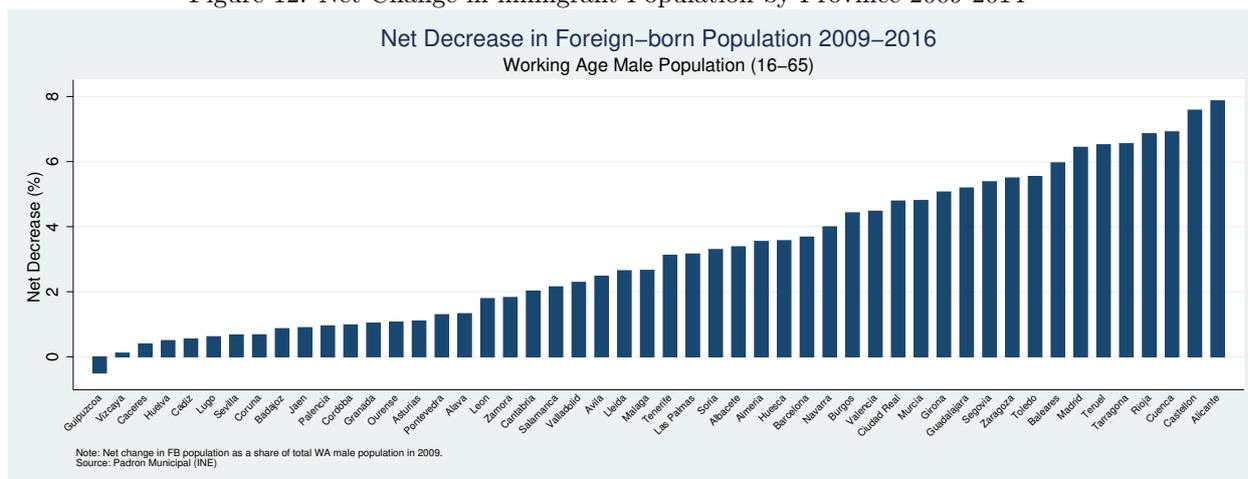
that they do not cause any empirical problem for my results.

⁵⁸There are also moves that take place within the province borders. My measure do not account for them as it captures moves internal moves which cross the provincial borders.

⁵⁹I consider an individual native if the person is born in Spain and has Spanish nationality.

⁶⁰In terms of immigrant population, these departures are even more striking. Figure 21 presents the same results as a share of the immigrant population in the province in 2009. The net departures between 2009 and 2016, correspond on average to 20% of the immigrant population in 2009.

Figure 12: Net Change in immigrant Population by Province 2009-2014



Net change in the total immigrant male working-age population between 2009 and 2014, as a share of the total working-age male population in 2009. Positive bars correspond to net outflow of the population. Data source: INE

Spanish Labor Force Survey

I supplement these data with the Spanish Labor Force Survey (EPA, or *Encuesta de la Población Activa* in Spanish). The EPA is conducted, every quarter, by the Spanish National Institute of Statistics (INE, or *Instituto Nacional de Estadística* in Spanish) with a sample of some 65,000 households (about 180,000 individuals) and is designed to be representative of the Spanish population at province level. I use the EPA for two reasons: first, I use it to calculate province specific rates (i.e. employment, participation rate, unemployment rates etc.). I prefer using the EPA as it includes employment both in formal and informal sector, while the MCVL, although much larger, is only representative of the formal employment. It is also preferable for estimating the unemployment rate as the MCVL records individuals while they are only receiving unemployment benefits, which results in an undercounting of the number of unemployed individuals.⁶¹ Second, although the *Padrón* is preferable for computing population due to its high precision, it lacks information on skill levels of the individuals. I use the EPA to compute population stocks by skill level for further analysis.

3.4.2 Summary Statistics

Table 1 presents the summary statistics. My sample covers 50 provinces, over the 2009-2014 period which leads to 250 province-year observations. The table presents the main variables in the regression. The

⁶¹As an administrative dataset, the MCVL records an individual as unemployed only if he is eligible for unemployment benefits. Individuals who have not accumulated enough contribution periods to be eligible for unemployment insurance or whose benefits have expired due to long unemployment spells do not appear in the data. This creates an undercounting in the actual number of unemployed individuals. Lafuente (2019) presents a simple and systematic method to expand the MCVL which makes it possible to correct for this undercounting, and obtain unemployment rates that match those obtained from the EPA.

first two lines of the table present average number of employed male native workers, and average salaries. I use these numbers to compute the annual growth rate in employment which I use as the dependent variable.

Table 1: Summary Statistics

Variable	N	Mean	sd	Min	Max
Native Workers	250	2,683.0	4,198.2	232.0	25,414.0
Monthly Wage, in Euros	250	1,668.2	215.1	1,339.8	2,283.8
Annual Decrease in FB	250	1,597.2	4,457.0	-2,437.0	36,309.0
Total Population	250	316,065.3	388,844.3	30,006.0	2,204,700.0

These are the main variables used in the analysis of the effect of out-migration on wage and employment growth of natives. The averages are unweighted, so do not necessarily coincide with the true average of Spain. The data covers the period 2009-2014 (5 years) and 50 provinces.

Data source: MCVL and INE

Table 1 also gives an idea about the annual out-migration that each province faced. During the period, Spanish provinces lost annually, on average, 1597 immigrant working-age male. Given that average working-age population in a province is around 316 000, this corresponded to an out-migration rate, or a reduction in the labor supply of 0.5% annually.

3.5 Empirical Strategy

In this section, I first explain how my main regression equations relate to the theoretical model presented in Section 3.2, and then describe my estimation and identification strategy.

3.5.1 Econometric Equation

I estimate the effect of net outflow of the immigrant population on the changes of labor market outcomes for native workers over the same period. Corresponding to my theoretical setup, I use the following first differences regression model:⁶²

$$\Delta \ln w_{g,j,t} = \beta_g \Delta \text{foreign}_{j,t} + \alpha_t + \varepsilon_j \quad (12)$$

$$\Delta L_{g,j,t} = \delta_g \Delta \text{foreign}_{j,t} + \alpha_t + \varepsilon_j \quad (13)$$

where

$$\Delta L_{g,j,t} = \frac{L_{j,t}^{\text{Native}} - L_{j,t-1}^{\text{Native}}}{L_{j,t-1}^{\text{Native}}} \quad \text{and} \quad \Delta \text{foreign}_{j,t} = \frac{\Delta N_{j,t-1}^{\text{foreign}} - N_{j,t}^{\text{foreign}}}{N_{j,t-1}^{\text{Native}} + N_{j,t-1}^{\text{foreign}}}$$

$\Delta \ln w_{g,j,t}$ is the change in mean of log wages of natives⁶³, in group (i.e, skill, sex, age, contract type) g , and province j , between two periods, $t - 1$ and t . $\Delta L_{g,j,t}$ is the percentage change in the

⁶²For other papers using first differences models on evaluating the impact of immigration, see Gonzalez and Ortega (2013); Ortega and Verdugo (2016); Dustmann et al. (2017a); Sanchis-Guarner (2017)

⁶³Some papers in the literature use changes in the log of mean wages. See Borjas et al. (2012) for a discussion on why this is an error.

native employment in group g , and province j , between two periods, $t - 1$ and t , $\Delta foreign_{j,t}$ is the net-change in immigrant population between two periods, divided by the province initial population, α_t is the time-fixed effect. Finally $\varepsilon_{j,t}$ is the random error term. Equations 12 and 13 are written in first differences to eliminate time-constant area and, when applicable, group fixed effects.

The variable of interest is the normalised net-change in immigrant population in province j (during a given period) divided by the province initial population.⁶⁴ The net-change is calculated as the difference between immigrant population between t and $t - 1$. Using the normalised net-change instead of (log) net change as the measure of net outflows eliminates any unobservables that might equally affect both the numerator (foreign-born outflows) and the denominator (original province population, sum of natives and immigrant). Standardising net-change by initial population stocks also deals with the fact that regions have different population and labor market dynamics (Card, 2001; Peri et al., 2009; Dustmann et al., 2017a; Lee et al., 2017). Moreover, scale effects can induce spurious correlation between higher outflows and higher changes in labor market outcomes. This correlation could arise due to the fact that the average and standard deviation of both variables are likely to be proportional to the population in the province.⁶⁵

Finally, consistent with the model, I use only the *total* but not the group-specific outflow of immigrant population. This approach is preferable as it does not require pre-allocation of immigrant to skill groups based on their observable characteristics, thus avoiding the problem of misclassification that arises when such observable characteristics are used to assign immigrants into skill groups in which they do not compete with natives.⁶⁶ It thus gives the *total* wage and employment effects of a decrease in labor supply due to out-migration of the immigrant population as derived in equation 11.

The parameters β_g and δ_g , measure the impact of the total net outflow of immigrant on the percent change in wages and employment of native workers in skill group g in area j between the two time periods. If wages are fully flexible, these parameters correspond to the expression derived in equations 6 and 11. If wages are partially rigid, the wage response β_g is determined exogenously by the degree of rigidity and employment response δ_g is given by equation 4. The employment response in equation 13 captures, employment movements across areas in addition to movements from and to non-employment (in activity or unemployment).⁶⁷

⁶⁴I compute these rates using only working-age male population, i) for consistency with the outcome treatment which includes only working-age male natives, ii) to better capture the changes in the labor market, by reducing the noise in the population moves which are not motivated by labor market conditions. In the following sections, I test the robustness of my results by using measures which are calculated using both female and male population, as well as all age groups.

⁶⁵Measuring changes in immigrants labor supply as a share of initial total population is standard in the literature (Dustmann et al., 2017a; Lee et al., 2017). Still, in section 3.7.4 I test the robustness of my results by using alternative measures. Note that I measure the out-migration rate as the difference in levels between $t - 1$ and t so that net outflows would have positive values, which makes the interpretation of the coefficients more intuitive.

⁶⁶Dustmann et al., 2013 show that immigrants often downgrade upon arrival, which Fernandez and Ortega (2008) show to be the case in the Spanish context as well. Thus assigning immigrants to skill groups based on observed characteristics may lead to serious misclassification. This estimation strategy is similar to Altonji and Card (1991); Dustmann et al. (2013, 2017a); Ortega and Verdugo (2016).

⁶⁷This differs from other studies which use the change in the local employment-to-population ratio such as Altonji and Card (1991); Boustan et al. (2010).

Finally, I estimate equation 12 and 13 by weighting the number of observation used to compute the dependent variables in each province-skill cell at base period and cluster the errors by province to account for potential location-specific correlations (Moulton, 1990).⁶⁸

3.5.2 Identification

My identification relies on exploiting the variation in the net outflow rate which, after controlling for province and time fixed effects, is uncorrelated with local determinants of labor market demand and economic performance between 2009-2014.⁶⁹ There are two issues regarding the identification. First, the immigrant population located in provinces which are more severely affected by the crisis will be more likely to leave. Second, the distribution of the immigrant across provinces before the crisis may not be random.

Concern 1: Out-migration and OLS as Lower Bound Estimates

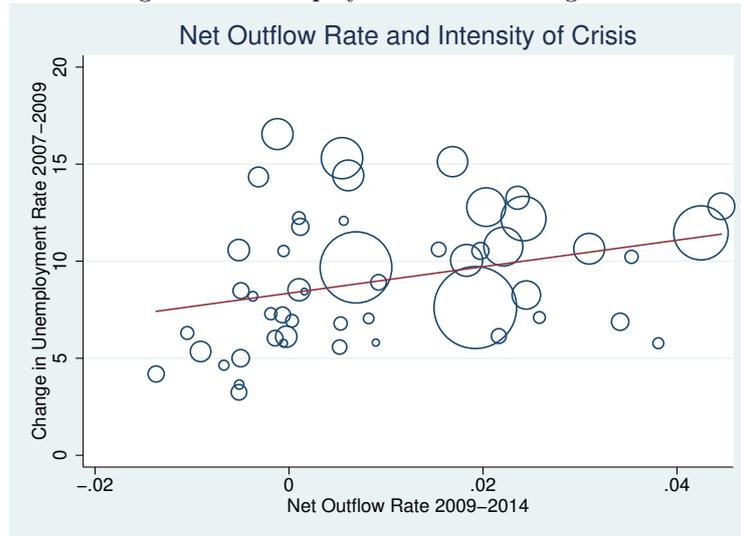
A first order concern is identifying a source of variation for the out-migration rate that is uncorrelated with local determinants of labor market demand and economic performance during the period. There is geographical variation in the intensity of the Great Recession across provinces. Figure 13 shows that provinces which experienced stronger increase in the unemployment rate between 2007-2009 also saw higher departures.⁷⁰

⁶⁸Similar papers use weights proportional to the number of observations used for the computation of the LHS (see for instance, Ottaviano and Peri (2012); Dustmann et al. (2017a); Lee et al. (2017)). Some papers such as Hunt (1992); Clemens and Hunt (2017) use the inverse of the sampling variance as weights. All the results presented in the present paper are robust to the use of both weights.

⁶⁹Although not presented, I test the endogeneity of net outflow rate by using an augmented regression tests (Durbin-Wu-Hausman test), as suggested by Davidson and MacKinnon (1993). The results show that OLS is not consistent and use of instruments is necessary. Results are available upon request.

⁷⁰I use the increase in the unemployment rate between 2007 and 2009 for two reasons. First, the dramatic and sudden increase in the unemployment rates happened during this period. In that sense, this measure captures well the depth of the crisis in each province. Second, I prefer using the increase in this period, as opposed to longer periods, as they are not dampened by the outflows that took place starting 2009 and thus reflect the initial demand shock. See Yagan (2019) for a similar approach. The correlation presented in the figure is robust to using longer windows.

Figure 13: Unemployment and Out-migration



The figure plots increase in the unemployment rate between 2007 and 2009 (y-axis) and Net Outflow Rate of the immigrant population between 2009 and 2014 (x-axis). Each point corresponds to one of the 50 provinces, excluding Melilla and Ceuta. Each observation is weighted by province immigrant population in 2009. Data source: INE

This creates a correlation between intensity of the demand shock and the out-migration rates.⁷¹ As out-migration is higher in provinces that experience stronger negative wage and employment shocks, this association induces a spurious negative correlation between out-migration and employment/wage growth that could lead to a negative bias in the OLS estimate. Hence, given the positive effects I estimate, the OLS estimator provides a lower bound for the actual effect of out-migration on mean wages and employment (see Dustmann et al. (2015), for a similar argument). In Section 3.7, I further address this issue by controlling for changes in the local demand.

Concern 2: Non-random Distribution of Immigrant Population and Endogenous Location of Native

The departure of immigrants is only possible if there is an immigrant population in the province in the first place. This initial distribution of the immigrants across provinces, however, may not be random. The use of first-differences takes care of province characteristics that are fixed over time which allows me to make progress towards the identification of β_g and δ_g .⁷² Still, the unbiased identification of these parameters requires the out-migration rate to be uncorrelated with the time-varying component of the error term. There is no prior on the direction of the bias as I am looking at the impact of outflows during a demand shock which is less straightforward than immigration in good times. For instance, if

⁷¹Although the negative demand shock and out-migration rates are correlated, there is still important spatial variations in distribution of immigrants in 2009, share of construction sector prior to the crisis and intensity of the demand shock as can be seen in the maps in Appendix Section 23.

⁷²The normalisation of the independent variable takes care of the concerns due to size of the immigrant population compared to the local population.

immigrants located in areas with faster wage and employment growth (conditional on all the time-varying and time-invariant controls) were more likely to leave, the estimated parameters would be upward biased. If, on the contrary, immigrants located in areas with slow growth rates were more likely to leave, then the parameters would be downward biased.

To deal with this identification issue, I construct an instrument adapting the “shift-share” methodology which is widely used in the literature.⁷³ Immigrants tend to locate disproportionately in areas where other immigrants from the same nationality or ethnicity have located in the past, to benefit from social and economic networks established by those who arrived earlier. I exploit this “past settlement instrument” and use the past spatial distribution of the immigrants in order to predict the current location patterns. Specifically, year-to-year variation of the national stocks (the “shift”) of different nationalities are distributed across provinces according to some historical distribution of immigrants (the “share”).

To construct the instrument, I first calculate the share of immigrants located in province j in 1991.

$$share_{j,1991}^n = \frac{Foreign_{j,1991}^n}{\sum_r^R Foreign_{j,1991}^n} \quad (14)$$

To obtain yearly predictions of the number of immigrants by nationality n for province j in year t , I multiply the expression 14 by annual national stock of immigrants $Foreign_{j,t}^n$ of nationality n . This stock is calculated adding the number of immigrants of that nationality in all provinces in Spain, in year t .⁷⁴ I leave-out the stocks in the same province, to address concerns that the introduction of own-area stocks may mechanically increase the predictive power of the instrument (Autor and Duggan (2003); Goldsmith-Pinkham et al. (2018)). The imputed immigrant stock of a specific nationality n in province j at time t is thus calculated allocating yearly total national stocks weighted by its historical share (14):

$$\widehat{FB}_{j,t}^n = (Foreign_{Spain,t}^n) * share_{j,1991}^n \quad (15)$$

To calculate the imputed total (all nationalities) immigrant stock in province j at time t , I sum (15) across nationalities (N):

$$\widehat{FB}_{j,t} = \sum_n^N (\widehat{FB}_{j,t}^n) \quad (16)$$

Note that the instrument is constructed from combining nationality-specific predictions for every province (a weighted sum of the national-minus-province inflows using the distribution of nationality in 1991 as weights). These predicted stocks generate a variation by exploiting differences in national flows and the initial distribution across labor markets which are arguably less endogenous to local economic

⁷³This strategy has been set by Altonji and Card (1991) and modified by Card (2009), and followed by many including Card (2001); Ottaviano and Peri (2006); Basso et al. (2018), applied in the case of Spain by Gonzalez and Ortega (2011, 2013); Sanchis-Guarner (2017); Fernández-Huertas Moraga et al. (2017); Fernández-Huertas Moraga (2017).

⁷⁴ $Foreign_{j,t}^n = \sum_{r \neq j}^R Foreign_{r,t}^n$

conditions.

As a final step, I calculate the change in predicted immigrant stocks and divided it by the imputed population (imputed immigrant plus native stock) in province i at the beginning of the period $t - 1$. The instrument is constructed as follows:

$$\Delta \widehat{foreign}_{j,t} = \frac{\widehat{FB}_{t-1} - \widehat{FB}_t}{\widehat{FB}_{j,t-1} + Native_{j,t-1}} = \frac{\Delta \widehat{FB}_{j,t}}{\widehat{Pop}_{j,t-1}} \quad (17)$$

For this instrument to be valid it has to be sufficiently correlated with the out-migration rate but uncorrelated with the local shocks that affect variations in the labor market outcomes of the natives, conditional on province and time fixed effects.⁷⁵ The relevance of the instrument can be assessed by the value of the F-statistics of the instrument in the first stage of the 2-stage-least-squares (2SLS) regressions, and additionally by using weak identification tests.⁷⁶

The validity of the instrument relies on the two components of Equation 15 being uncorrelated with local shocks that affect outcome variables. Regarding the local share of immigrants by nationality in the base year, the exclusion restriction requires that the only channel through which immigrant geographical distribution in 1991 affects current changes in labor market outcomes is through its influence on shaping the current immigrants location patterns, conditional on fixed effects. In other words, the unobserved factors determining the location of immigrants in one province with respect to another in the base year (1991) have to be uncorrelated with the relative economic prospects of the provinces during the period of analysis (2009-2014). I consider 1991 to be separate enough from 2009-2014 for immigrant shares to be uncorrelated with the past demand shocks.⁷⁷ Still, it is possible that unobservable shocks correlated with local labor market conditions that affected immigrants' location decisions in 1991, continued in the following periods. In Appendix Section 3.13.12, I show factors that were determinant in the distribution of the immigrant population in 1991 were not relevant in the distribution during the period of analysis.

Furthermore, as shown by Jaeger et al. (2018) for the US, such instrument can be problematic if the location choice of immigrants and country-of-origin mix are stable over time.⁷⁸ In the context of Spain these problems are likely to be milder as the local immigrant inflows over time are much less correlated as both country-origin mix and destinations have changed greatly for the immigrants between 1991 and

⁷⁵Taking the first differences addresses concerns about the differences due to province characteristics which are fixed over time.

⁷⁶All the result tables in section 3.6 provide the Kleibergen-Paap statistics (test of weak identification), which is robust to non-i.i.d error terms, and corresponds roughly to the t-stat of the included instruments in the first-stage to the square.

⁷⁷Spain went through an important economic crisis (1992-1993) followed by economic recovery and growth (from 1997). Given the changes in the economy, it is unlikely that 1991 immigrants were able to predict these future shocks (or any other shock not captured in the area/time fixed effects) 18 years before my period of analysis starts. Use of past settlement patterns that are sufficiently lagged are important for validity of this instrument. For such an argument see for instance Dustmann et al. (2013); Orrenius and Zavodny (2010). In Section 3.7.3 I check the robustness of my results using earlier and later base years.

⁷⁸Jaeger et al. (2018) show that if the same locations keep receiving immigrant inflows, firms located in cities receiving large waves of immigrants would progressively raise their capital stock, which would push up the wages in these cities relative to others. As the same cities repeatedly receive the inflows, the adjustment process of past waves (of arrivals) overlaps with that of the new waves. This "overlapping response problem" makes it difficult empirically to separate the (presumably) negative short-run wage effect of immigration from the (potentially) null medium or long run wage response.

2009. For instance autocorrelation coefficients for both the observed immigrant shock $\Delta foreign_{j,t}$, and the shock predicted by my instrument $\Delta \widehat{foreign}_{j,t}$, is 0.55. This correlation level is dramatically lower than those observed in the U.S. by Jaeger et al. (2018) where they are above 0.9 since the 1990s. Regarding the country-origin mix, the serial correlation in national composition between 1991 and 2009 is 0.67. If I exclude the Moroccans, which constituted the biggest immigrant group in 1991 and the second biggest in 2009, the correlation drops to 0.39. Both of these numbers are much lower than levels found by Jaeger et al. (2018) for the US which are between 0.9-0.99 starting from 1970s. Both figures suggest that the “Overlapping Response Problem” seem to not be of an important issue in the context of Spain. Still, in order to alleviate remaining concerns, in Appendix Section 3.6.1, I carry out the “Multiple Instrumentation” procedure as suggested by Jaeger et al. (2018) and show that the results are robust to inclusion of lagged outflows.

A final issue with the construction of the instrument is the endogenous location choices of natives as a response to immigrant mobility. The total population stock, which appears in the denominator, is the result of the sum of the (imputed) immigrant plus the native’s stocks. The number of total natives residing in a given province might depend on the number of immigrant in the same location or on unobservables correlated with the labor market outcomes. For this reason, I use a similar shift-share strategy to predict the location choice of natives, based on past location patterns similar to the immigrant and replace the actual native stock by its imputed number. Details of this procedure are given in the Appendix Section 3.13.11.

I use this instrument in my estimation and different versions of it in the robustness checks. In Section 3.6.1 I discuss and test the validity of this instrumental variable approach, and in Section 3.7.3 I check if the results are robust to using different definitions of the shift and share in the construction of the instrument.

3.6 Main Results

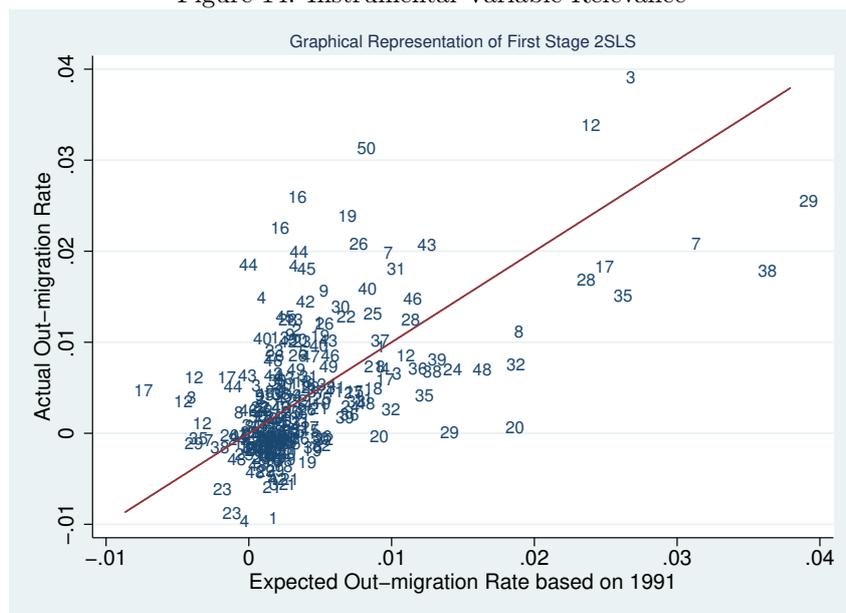
In this section, I present the first stage results and discuss the validity of the instrument, and then present the second stage results.

3.6.1 First Stage and Validity of Instruments

My main identification strategy consists in implementing the 2SLS estimation outlined in Section 3.5.2.

Figure 14 provides a visual representation of the first stage. It plots the predicted out-migration rate in the horizontal axis against the actual out-migration rate. Each observation in the figure corresponds to a province-year observation between 2009-2014. The figure shows that actual and predicted out-migration rates are strongly correlated.

Figure 14: Instrumental Variable Relevance



The figure plots the predicted out-migration rate in each province-year pair between 2009 and 2014 (based on the distribution in 1991) in the horizontal axis against the actual out-migration rate between 2009 and 2014. The figure represents a graphical visualization of the first stage of the two-stages-least-squares estimation. Data source: INE

In Table 2 I show the coefficients from the first-stage regression:

$$\Delta foreign_{j,t} = \theta \Delta \widehat{foreign}_{j,t} + \alpha_t + \varepsilon_{j,t}$$

The coefficient θ , reported in the first row of the table, represents the effect of the imputed change in immigrant population (obtained as in Equation 17) on the actual change in immigrant as a share of the total working age population, which is the explanatory variable in my second-stage regression. All of the regressions include time-fixed effects and errors are clustered at province-level.

Table 2: First Stage Regressions, 2009-2014

	2	3	4	5	6
Predicted Net Outflow Rate	0.4145 (0.099)***	0.3769 (0.097)***	0.3150 (0.057)***	0.2300 (0.065)***	0.2422 (0.060)***
Compositional	No	Yes	Yes	Yes	Yes
Pop Growth	No	No	Yes	Yes	Yes
Economic 2009	No	No	No	Yes	Yes
Bartik	No	No	No	No	Yes
N	250	250	250	250	250
r2	0.53	0.55	0.74	0.78	0.78
Cragg-Donalds Stat	48.20	25.93	114.60	123.36	145.43

The table reports first-stage results. The dependent variable is the change in immigrant population between year t and $t - 1$, relative to total working-age population in year $t - 1$. The explanatory variable is the imputed change in immigrant population during the same period. The unit of analysis is province. Weights correspond to the number of employed natives in the base period ($t - 1$). Compositional characteristics correspond to changes in average age and schooling of the native workers. Population Growth corresponds to change in the total population between year t and $t - 1$, relative to total working-age population in year $t - 1$. Characteristics of the Economy in 2009 include share of agriculture, manufacturing, construction and services in total employment. Bartik refers to the change in predicted total employment. Regressions include year-fixed effects. Standard errors are clustered at province-level. Data source: MCVL and INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In the first column of Table 2, I report the regression coefficient weighted by the total working age population in the province in the base period. The estimated coefficient is highly significant and around 0.41. Specifically an increase in the imputed out-migration rate by one percentage point leads to a 0.41 percentage point increase in the actual out-migration rate between 2009-2014.

In my baseline specification which is in first differences, I only include the out-migration rate and time-fixed effects. However, in Section 3.7.2 I test the robustness of my results by adding various controls in order to address concerns about spurious correlations between the variable of interest and the outcome variables. In columns 2-5, I add these controls (which I explain in detail in Section 3.7.2) gradually.

Across all specifications, the imputed out-migration rate remains stable and highly significant. In all cases, the F-statistic remains high showing that the instrument is strong.

3.6.2 Second Stage

Table 3 presents the results of the estimation of equation 12 and 13. As explained previously, the parameters β_g and δ_g correspond to the total effect and they capture the combined impact of changes in labor market outcomes of the natives due to out-migration and native mobility. These results are obtained using data on annual changes on average wages and employment of natives during the period 2009-2014. In all specifications I include time dummies to control for national shocks, and cluster the standard errors at the province level. Each regression is based on a sample of 250 observations, for 50 provinces and 5 time periods.

Table 3: Effects on Wages and Employment of Natives, 2009-2014

	Wage		Employment	
	OLS	2SLS	OLS	2SLS
Net Outflow Rate	0.2092 (0.248)	2.0286 (0.609)***	1.1586 (0.348)***	2.4122 (0.752)***
N	250	250	250	250
Adj R2	0.49	0.26	0.53	0.49
KP F-Stat		17.45		17.45

The table reports OLS and 2SLS estimates for the impact of net change in immigrant population in an area, measured as the decrease in the number of working-age male immigrant population between t and $t-1$ as of total working-age male population in year $t-1$, on native local wage and employment growth in the aggregate. Columns 1 and 2 report OLS results, while columns 3 and 4 report 2SLS results. Regressions are estimated annually, across 50 provinces over 5 periods. Regressions are weighted by total employment in the base year and include year-fixed effects. Standard errors are clustered at province-level. Data source: MCVL and INE.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The first and third columns of Table 3 show the OLS results obtained for wage and employment growth, respectively. Remember that I do not include province dummies as the analysis is in first differences. This simple correlation is 0.2 for wages and 1.15 for employment. Both models have high explanatory power (the R^2 around 49% for wages and 53% for employment). In order to make causal claims about the estimates, I implement the IV strategy explained in section 3.5.2.

In columns 2 and 4, I repeat the exercise by instrumenting the net outflow rate using the *predicted* net outflow rate based on the distribution of the immigrant population in 1991. Similar to the OLS results standard errors are clustered at the province level. The table displays the test of the strength of the instruments (F-stat Kleibergen-Paap). As expected, in all specifications the standard errors increase when using instrumental variables.

My instrumental variable approach confirms the spurious correlation between the depth of the recession and the outflow rates. As explained in section 3.5.2, provinces that were hit harder by the recession saw higher outflow rates. Introduction of the instrument takes care of this downward bias and increases the estimated coefficients. In all models the instrument is strong and the Kleibergen-Paap F-stat is above the Stock-Yogo critical values. Over the 2009-2014 period, I find that the outflow of the immigrant population has positive effects on the wages and employment of native workers. The estimates imply that a 1 percentage point increase in the annual outflow rate increases the wages and employment in the province by 2% and 2.4% respectively.

3.7 Robustness

In this section, I carry out various exercises to test the robustness of these results. I start by examining whether pre-trends or other spurious correlations drive the results. Later, I show that the results are

robust to the use of different outflow measures, alternative instruments, and weights.

3.7.1 Previous-trends

It is important to check that observed changes in the wages and employment are not driven by persistent unobserved factors and previous trends. In order to explore this I carry out three tests.

First I regress pre-recession changes in wages and employment on the total outflow rate from 2009-2014. More specifically I use the total out-migration rate over the 5 year period, and test whether it has any explanatory power over the changes in the pre-crisis period. Similarly to the main specification, I weight the regressions with employed population in the base year (i.e. 2009). Table 4 shows this falsification test for the period 2003-2008.⁷⁹ The OLS and 2SLS results for wages are presented in columns 1 and 2, while results for the employment are in columns 3 and 4. I find a negative and non-significant relationship for wages, and negative and highly significant results for employment. These results imply that, if anything, the provinces which have witnessed higher outflow rates during the Great Recession were in a negative wage and employment growth trend before the Recession.

Table 4: Falsification Test 1, Growth in 2003-2008

	Wage		Employment	
	OLS	2SLS	OLS	2SLS
Net Outflow Rate 2009-14	-0.0179 (0.019)	0.0017 (0.035)	-1.8976 (0.701)***	-3.2787 (1.671)**
N	50	50	50	50
KP F-Stat		23.13		23.13

The table reports OLS and 2SLS estimates for the impact of net change in immigrant population in an area, measured as the decrease in the number of working-age male immigrant population between 2009 and 2014 as of total working-age male population in year 2009, on native local wage and employment growth of natives in the aggregate in the previous period. Columns 1 and 2 report results on wage growth, while columns 3 and 4 report results for employment growth. Regressions are weighted by total employment in each province in 2009. Standard errors are clustered at province-level.

Data source: MCVL and INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Second, in the spirit of Dix-Carneiro and Kovak (2017), I regress equations 12 and 13, but include the pre-recessions changes in outcomes ($y_{j,2008} - y_{j,2003}$) to address the possibility of confounding pre-existing trends. Table 5 presents the 2SLS results. Columns 1 and 3 present baseline results for comparison. In columns 2 and 4, I add the pretrends for wages and employment that were calculated as the change in the outcomes from 2003 to 2008. Pretrends for both wages and employment are small and statistically insignificant. These results confirm the findings above.

⁷⁹I do this exercise by looking at the change in the 5 year period prior to the crisis as it corresponds to the length of my analysis period (i.e., 2009-2014). The results are robust to changing the start and end years in the pre-crisis period, or testing for longer/shorter differences.

Table 5: Falsification Test 2 , Pre-trends

	Wage		Employment	
	2SLS	2SLS	2SLS	2SLS
Net Outflow Rate	2.0286 (0.609)***	2.0680 (0.724)***	2.4122 (0.752)***	1.8584 (0.417)***
Pre-trend wage		0.0188 (0.089)		
Pre-trend employment				-0.0485 (0.031)
N	250	250	250	250
KP F-Stat	17.45	16.02	17.45	34.78

The table reports 2SLS estimates for the impact of net change in immigrant population in an area, measured as the decrease in the number of working-age male immigrant population between 2009 and 2014 as of total working-age male population in year 2009, on native local wage and employment growth of natives in the aggregate (skilled and unskilled) in the previous period. Columns 1 and 2 report results on wage growth, while columns 3 and 4 report results for employment growth. The unit of observations are provinces. Regressions are weighted by total employment in 2009 and include year-fixed effects. Standard errors are clustered at province-level.

Data source: MCVL and INE.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Third, following Lee et al. (2017) I test the correlation between the share of immigrant in the 2009 population and the pre-trends for changes in wages and employment. I do this test in order to check whether distribution of immigrant before the crisis is correlated with the outcomes in the following period. I test for two different periods, 1998-2008 and 2003-2008. Results in Table 6 show that the immigrant distribution in 2009 is not correlated with the wage growth of natives in the previous period. It is however negatively correlated with the employment growth of natives. This shows that, provinces with higher share of immigrant population in 2009 were in negative employment growth trend for natives, which is in line with the results in the previous falsification tests.

Table 6: Falsification Test 3 , Share of immigrant

	Wage		Employment	
	1998-2008	2003-2008	1998-2008	2003-2008
Foreign Share in 2009	0.0024 (0.015)	-0.0030 (0.005)	-6.5886 (2.364)***	-0.6329 (0.166)***
N	50	50	50	50
Adj R2	0.00	0.01	0.19	0.31

The employment and wage growth in the previous periods, standardized by initial employment. The explanatory variable is the share of immigrant in the total working age population in 2009. The unit of observations are provinces. Regressions are weighted by total employment in 2009. Standard errors are clustered at province-level.

Data source: MCVL and INE.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

These three exercises show that province-specific pre-trends did not drive the positive wage and employment growth observed in provinces with higher immigrant outflows. On the opposite, native

wages and employment in provinces with higher outflows during the crisis were in a negative growth trend before the crisis.

3.7.2 Bartik and Other Controls

The effects found in the previous section show that departure of the immigrant population improved the wages and the employment outcomes of the staying natives. Could these results be driven by factors that are correlated with the outflows? In order to check the robustness of these results, I test the specification while including controls that may be correlated with the evolution of the labor market outcomes during the period.

Table 7 presents results where I include various controls. columns 1 and 4 present 2SLS results without any controls, as a benchmark. In columns 2 and 6, I control for changes in demographic characteristics, namely changes in average age and share of high-skilled workers among the total employed between two periods.⁸⁰ If these variables affect economic performance during this period, their inclusion reduces the risk of spurious correlation. Inclusion of these variables do not change the significance of my variable of interest and my instrument remains strong.

⁸⁰As pointed out in Verdugo (2016), during the Recession the labor force became older and more skilled as the younger and less-skilled workers were first to be fired. Inclusion of these controls captures increases in the wages driven by compositional changes.

Table 7: Wages and Employment with Controls

	Wage			Employment				
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS		
Net Outflow Rate	2.0286 (0.609)***	2.0066 (0.674)***	2.4734 (0.966)**	2.8256 (1.147)**	2.4122 (0.752)***	2.7436 (0.829)***	2.6549 (0.965)***	2.8845 (1.089)***
Demographic	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Economic 2009	No	No	Yes	Yes	No	No	Yes	Yes
Bartik	No	No	Yes	Yes	No	No	Yes	Yes
Native Outflows	No	No	No	Yes	No	No	No	Yes
N	250	250	250	250	250	250	250	250
KP F-Stat	17.45	14.99	9.67	7.97	17.45	14.99	9.67	7.97

The table reports 2SLS estimates for the impact of net change in immigrant population in an area, measured as the annual decrease in the number of working-age male immigrant population between period t and $t - 1$ as of total working-age male population in t , on native local wage and employment growth. The unit of observations are provinces. Regressions are weighted by total employment in period t . Characteristics of economy in 2009 include share of agriculture, manufacturing, construction and services in total employment. Bartik refers to the change in predicted total employment. Native out-flow is measured as the annual decrease in the number of working-age male native-born population between period t and $t - 1$ as of total working-age male population in t . Regressions are weighted with total employment in period $t - 1$ and include year-fixed effects. Standard errors are clustered at province-level.

Data source: MCVL and INE.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The economic structure at the start of the period (2009) may also be correlated with the presence of the immigrant population and the labor market performance of each province throughout the period. In order to control for differences in the structure of economic activity, in columns 3 and 7 I include the share of agriculture, construction, and services in the total employment.⁸¹ As described earlier, although the immigrants were employed in agriculture and services, they were also highly concentrated in the construction sector.

More importantly, the period under study is that of an economic crisis. The severity of the Great Recession differed significantly across provinces. This period saw a significant decline in economic activity in several specific industries. The sectoral composition of provinces might explain a large part of the employment performance and could be correlated with the immigrant presence. In line with my model, I need to control for shifts in demand that is independent from local labor supply characteristics. Following the literature studying the local economies, I introduce a variant of “shift-share”, $\Delta Bartik_{j,t}$, shocks à la Bartik (1991). I detail its construction in Appendix Section 3.13.13.

One important remaining issue is the potential omitted variable bias that may result from native outflows. As discussed in 3.10.1, native mobility was minimal before and limited during the crisis period. Still, if natives’ out-migration rates follow a similar distribution to that of immigrants, the estimated parameters may be capturing the combined effects of both immigrant and native departure. As Appendix Figure 22 show, out-migration rates for immigrant and natives’ were not correlated. Still, to address this concern, in columns 4 and 8 I control for native mobility. I construct the native outflow rate as the net change in the native-born population between two periods, normalized by the total population in the base period, similar to the immigrant outflow rate.⁸² The results hold.⁸³

Overall results show that, the effects of out-migration rate is not due to spurious correlations caused by the economic and demographic conditions at the beginning of the crisis and the changes that followed during the period.⁸⁴

3.7.3 Alternative Instruments

In this section, I test the validity of my results using alternative instruments based on different shifts and shares. Table 8 presents the regression results where I use different instruments. Panel A presents results for wage growth while Panel B presents the results for employment growth.

The first column shows the baseline estimates, which uses the instrument where the share is the

⁸¹I use the sectoral shares in employment for the year 2009 as it is the base year of the analysis. Use of base year characteristics is common practice in similar exercises (see for instance Boustan et al. (2010); Autor et al. (2013); Lee et al. (2017)). The results also hold if I use the shares in 2007.

⁸²Native outflows can suffer from measurement error as Spanish citizens are not deleted after two years. In Appendix Section 3.13.10, I discuss this issue and explain why it does not any threat to the empirical exercise.

⁸³In models with the full set of controls, the F-test is below the threshold value of 10. If I use alternative instruments (which I present in the next section), I find similar elasticities and the instruments pass the weak instrument test.

⁸⁴Throughout the paper, I present results with only out-migration rate, without including the controls. I prefer not including the controls for transparency. All of the results presented in the paper are however robust to inclusion of these controls.

provincial distribution of the immigrant in 1991 and both the immigrant and native stocks in the denominator are predicted to account for endogenous allocation. In column 2, I use the predicted immigrant stocks based on the distribution in 1991, and the actual stock of natives in the denominator (as opposed to predicted native stocks in the baseline). In column 3, I use as shifts nationality stocks predicted through a gravity model, to address issues related to exogeneity of the national stocks. Details of the gravity procedure are given in the Appendix 3.13.11. In column 4, I distribute population stocks based on country of birth (as opposed to nationality).

In the following columns, I construct alternative instruments based on the distribution of immigrant population in 1996 (columns 5 and 6), in 1999 (columns 7 and 8) and in 2001 (columns 9 and 10). For each base year, I present results where I use the actual yearly stocks as shift (columns 1, 2, 4, 5,7 and 9) or stocks predicted through the gravity model (columns 3, 6, 8 and 10).

Table 8: Alternative Instruments

Panel A: Change Wages											
	Baseline	1991	Gravity 91	1991 FB	1995	Gravity 95	1999	Gravity 99	2001	Gravity 2001	
Net Outflow Rate	2.0286 (0.609)***	2.3059 (0.678)***	1.9862 (0.971)**	2.2348 (0.722)***	1.7510 (0.541)***	1.2842 (0.884)	1.5724 (0.522)***	1.4740 (0.866)*	1.0621 (0.436)**	1.1566 (0.732)	
N	250	250	250	250	250	250	250	250	250	250	
KP F-Stat	17.45	15.76	9.00	13.69	36.50	17.66	47.13	17.39	93.08	19.68	
Panel B: Change Employment											
	Baseline	1991	Gravity 91	1991 FB	1995	Gravity 95	1999	Gravity 99	2001	Gravity 2001	
Net Outflow Rate	2.4122 (0.752)***	2.3127 (0.786)***	1.8147 (1.077)*	2.4424 (0.875)***	2.0853 (0.453)***	2.0018 (0.809)**	1.9260 (0.364)***	1.9498 (0.824)**	1.7013 (0.321)***	2.1853 (0.777)***	
N	250	250	250	250	250	250	250	250	250	250	
KP F-Stat	17.45	15.76	9.00	13.69	36.50	17.66	47.13	17.39	93.08	19.68	

The table reports 2SLS estimates for the impact of net change in immigrant population in an area (see the text for details on the measurement), on native local wage and employment growth of natives. The unit of observations are provinces. Regressions are weighted by total employment in year $t - 1$, and include year fixed-effects. Standard errors are clustered at province-level.
Data source: MCVL and INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All of the instruments, regardless of the shift or share that is used, pass the weak instrument test. As expected, the F-statistic gets larger when I use a share from a later period, while the use of national stocks predicted through gravity reduce the strength of the instruments. Although elasticities vary depending on the instrument, they are not statistically different from each other and my baseline results. These results confirm that the results are not dependent on the precise share and shift that I use in constructing the instrument.

3.7.4 Alternative Measures of the Supply Shock

The baseline measure of the local supply shock induced by the outflow of the immigrant population is $\Delta foreign_{j,t}$, which is the immigrant departures as a share of the total population in the base period. Papers that are focusing on the impact of immigration use various definitions for supply shock. In this section, I reconstruct the outflow rate following different measures used in the literature to show that my results are robust to alternative shock measures.

Table 9 reports the IV estimates of the parameters β_g and δ_g for various specifications using the same instrument, as in section 3.6.

columns 1 and 5 provide results for the measure used earlier as a benchmark. In column 2 and 6, following Card and Peri (2016) I define the out-migration rate as $\Delta foreign^{Card\&Peri}_{t,j} = \Delta N_j^{Foreign} / N_{j,t-1}^{Natives}$, where $\Delta N_j^{Foreign}$ is the net change in the number of immigrant working-age population between time t and $t - 1$, $N_{j,t-1}^{Natives}$ is the number of working-age natives in the base period. In columns 3 and 6, I define the out-migration rate similar to Friedberg (2001) as $\Delta foreign^{Friedberg}_{t,j} = \Delta N_j^{Foreign} / N_{j,t}^{Natives}$, which uses the number of natives in the current year in the denominator. Finally, following Hunt (1992) I define the out-migration rate as $\Delta foreign^{Hunt}_{t,j} = \Delta N_j^{Foreign} / N_{j,t}^{Natives} + N_{j,t}^{Foreign-born}$, where the denominator is the total working-age population in period t .

The results in Table 9 support the findings in section 3.6, and show that regardless of how out-migration rate is defined, the outflow of the immigrant increased both the wages and employment of natives.

Table 9: Alternative Measures

	Wage				Employment			
	Baseline	Card(2016)	Friedberg(2001)	Hunt(1992)	Baseline	Card(2016)	Friedberg(2001)	Hunt(1992)
Net Outflow Rate	2.0286 (0.609)***				2.0853 (0.453)***			
Foreign/Natives in t-1		1.7560 (0.771)**				2.0912 (0.603)***		
Foreign/Natives in t			1.7511 (0.768)**				2.0854 (0.602)***	
Foreign/Total in t				2.4955 (1.137)**				2.9719 (0.930)***
N	250	250	250	250	250	250	250	250
KP F-Stat	17.45	13.63	13.58	11.58	36.50	13.63	13.58	11.58

The table reports 2SLS estimates for the impact of net change in immigrant population in an area (see the text for details on the measurement), on native local wage and employment growth of natives. The unit of observations are provinces. Regressions are weighted by total employment in 2009. Standard errors are clustered at province-level. Data source: MCVL and INE.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As explained earlier, I measure the out-migration rate only taking into account changes in the working-age *male* population. As a robustness in Table 10, I re-calculate the measure males of all age groups (columns 2 and 4) and working-age population including women (columns 3 and 6).⁸⁵ The results can be seen in Table 10.

Table 10: Alternative Measures

	Wage			Employment		
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Net Outflow Rate	2.0286 (0.609)***			2.4122 (0.752)***		
All Ages Foreign male		1.8001 (0.632)***			2.1405 (0.681)***	
Net Outflow Rate (including women)			2.3232 (0.674)***			1.9241 (0.738)***
N	250	250	250	250	250	250
KP F-Stat	17.45	20.35	20.87	17.45	20.35	20.87

The table reports 2SLS estimates for the impact annual net outflow of immigrants (see the text for details) on native's wage and employment growth. The unit of observations are provinces. Regressions are weighted by total employment in year $t - 1$. Standard errors are clustered at province-level. Data source: MCVL and INE.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As expected inclusion of all age groups decreases the estimated elasticity slightly. Inclusion of female immigrant population gives slightly larger point estimates for wages while giving smaller estimates for employment. Despite these differences, I cannot reject the null hypothesis that differences in the elasticities are statistically significant. Overall, these results confirm that those results are not driven by the numerator or denominator chosen for the construction of the outflow measure.

3.7.5 Alternative Weights

All of the results presented in section 3.6 use weights that are proportional to the number of observations used to compute the LHS variable as it corrects for heteroskedastic error terms and thereby achieve more precise estimation of coefficients (Solon et al. (2015)). Use of such weights naturally, changes the importance of each province-year observation as more populated provinces are assigned more weight. In order to test the importance of weights in the estimation, I run the main regressions with and without weights. Table 11 reports 2SLS results for both wages and employment. Results in column 1 and 2 are the baseline, where regressions are weighted by number of employed in the base period. In columns 3 and 4, there are no weights.

Three results stand out: First although the coefficients for wages hardly change, the coefficient for employment gets larger which suggest that growth in native employment was especially important in less populated provinces.⁸⁶ Second, weighting by the number of observations least to smaller standard

⁸⁵I reconstruct all the instruments to reflect the changes in the group.

⁸⁶It should be noted however that given the standard errors, one cannot reject the hypotheses that both coefficients are statistically different from each other.

errors and thus a more precise coefficient estimation.⁸⁷ Finally, both weighted and unweighted estimates are consistent with each other which indicates that the model provides a good approximation.

Table 11: Alternative Weights

	Employment Weights		No Weights	
	Wage	Emp.	Wage	Emp.
Net Outflow Rate	2.0286 (0.609)***	2.4122 (0.752)***	2.7422 (1.357)**	3.0470 (1.498)**
N	250	250	250	250
KP F-Stat	17.45	17.45	6.54	6.54

The table reports 2SLS estimates for the impact of net change in immigrant population in an area, measured as the decrease in the number of working-age male immigrant population between t and $t - 1$ as of total working-age male population in year $t - 1$, on native local wage and employment growth in the aggregate. Regressions are estimated yearly, across 50 provinces. Regressions in columns 1 and 2 are weighted by total employment in the base year. All results include year fixed effects. Standard errors are clustered at province-level. Data source: MCVL and INE.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.8 Dynamics of Adjustment

In this section, I explore the dynamics of the adjustments for wages and employment. The panel dimension of my data allows me to test the effects of out-migration in longer differences and provide estimates that are comparable with some of the literature which are in longer differences. In Table 12 in addition to my benchmark analysis (which is in one-year differences), I also look at longer time windows (i.e., in longer differences) to study the changes in the elasticities. Specifically, I look at the changes between 2009 and 2013 (4 - year difference) and between 2009 and 2016 (7 - year difference).

To do so, I reconstruct out-migration rates, the instruments, and the outcome variables in longer differences. I use the same specification as in the previous section with first differences (FD).⁸⁸ I limit the observations so that both the start and the end date of the period remain within the period of interest (i.e., 2009-2016).⁸⁹

Table 12 presents the results for wage and employment growth. Columns 1 and 4 present the baseline results (which are annual) for comparison. Columns 2 and 5 report the elasticities for the change between 2009 and 2013, while columns 3 and 6 report those for 2009 and 2016. It can be seen that while the positive effects of outflows on wages persist until the fourth year, and disappears afterward. The employment effects, on the other hand, remain even in a 7-year window.

⁸⁷As discussed in Solon et al. (2015), when the number of observations used to compute the LHS is highly variable and small in some, using weights can improve the precision.

⁸⁸I further check the robustness of the results by adding the same controls as in Section 3.7.2. Results can be provided if requested.

⁸⁹I make this condition to make sure that my estimation captures the variation that takes place during the crisis period. I can relax this condition by allowing all the observations with either the start or the end period falling between 2009-2016. The results do not change.

Table 12: Longer Differences: Employment

	Wage			Employment		
	Baseline	2009-2013	2009-2016	Baseline	2009-2013	2009-2016
Net Outflow 1-year differences	1.6195 (0.421)***			2.2611 (0.785)***		
Net Outflow. 4-year differences		1.1236 (0.268)***			3.7171 (0.874)***	
Net Outflow. 7-year differences			0.5119 (0.413)			2.5436 (0.432)***
N	350	50	50	350	50	50
KP F-Stat	15.73	34.82	73.41	15.73	21.41	73.41

The employment and wage growth in the previous periods, standardized by initial employment. The table reports 2SLS estimates for the impact of net change in immigrant population in an area, measured as the decrease in the number of working-age male immigrant population at different time differences, on native local wage and employment growth in the aggregate. The unit of observations are provinces. Regressions are weighted by total employment in 2009. Standard errors are clustered at province-level. Data source: MCVL and INE.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This exercise shows that the outflows have employment effects that are time-persistent. Positive wage effects on the other hand disappear after 4 years. This “recovery” is clearly faster than what is observed following an immigration episode, where the negative effects due to immigration disappear after 5 years in Monras (2019) or 14-20 years in Blanchard and Katz (1992); Jaeger et al. (2018); Edo (2017).

3.9 Heterogeneity Analysis

The model predicts (11) that the out-migration of the immigrant population would have a stronger impact on the wage and employment outcomes of natives that have higher substitutability with the departing population. However, depending on the groups’ supply elasticities or wage rigidities, the final effect may differ in magnitude or the margin (i.e., employment vs. wage). In this section, I test the heterogeneous effects of outflows across groups with different skills and demographics.

In all regressions, I use the exact same specification and regress group-specific changes in the outcomes on the out-migration rate of working-age immigrant males, using the same instrument. I cluster the errors by province but weight each regression with the number of observations used to compute the changes in the LHS, which is specific to the group that is analyzed.

3.9.1 By Skill Group

As discussed earlier, most of the departing immigrant were from the lower half of the education distribution. For Spain, this means they have completed primary school and may have some secondary school. The MCVL includes information on the highest education level attained by the worker. Using this information, I group workers into two skill groups, those who have less than a university degree (i.e.,

low-skilled) and those who have a university degree or more (i.e., high-skilled).⁹⁰

Panel A in Table 13 presents the OLS and 2SLS results for wage growth for different skill groups. The OLS results presented in columns 1 to 3, show that outflow rate's effect on wages are statistically insignificant which is due to the negative correlation between out-migration and economic shock that was discussed above. Once the bias is taken care of with the instrument, all the coefficients become much larger, positive and significant.⁹¹

Table 13: Wages and Employment by Skill Group

Panel A: Change Wages						
	OLS			2SLS		
	All	Low Skilled	High Skilled	All	Low Skilled	High Skilled
Net Outflow Rate	0.2092 (0.248)	0.1289 (0.182)	0.0227 (0.380)	2.0286 (0.609)***	1.5773 (0.470)***	1.7531 (1.114)
N	250	250	250	250	250	250
Adj R2	0.49	0.41	0.43	0.26	0.26	0.36
KP F-Stat				17.45	16.74	17.90
Panel B: Change Employment						
	OLS			2SLS		
	All	Low Skilled	High Skilled	All	Low Skilled	High Skilled
Net Outflow Rate	1.1586 (0.348)***	0.9453 (0.296)***	1.2907 (0.402)***	2.4122 (0.752)***	1.8168 (0.777)**	4.6534 (1.400)***
N	250	250	250	250	250	250
Adj R2	0.53	0.55	0.16	0.49	0.54	-0.13
KP F-Stat				17.45	16.74	17.90

The employment and wage growth in the previous periods, standardized by initial employment. The table reports OLS and 2SLS estimates for the impact of net change in immigrant population in an area, measured as the decrease in the number of working-age male immigrant population between t and $t - 1$ as of total working-age male population in year $t - 1$, on native local wage and employment growth in the aggregate. Regressions are weighted by group-specific employment in base period. Standard errors are clustered at province-level. Data source: MCVL and INE.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel B in Table 13 reports the OLS and 2SLS results for employment growth for the same skill groups. Both the OLS and 2SLS, results show the employment effects of the outflows of the immigrant benefited natives of all skill groups. Similar to the wage results, 2SLS coefficients are much higher than the OLS ones as they take care of the downward bias. The results show that a 1% increase in the out-migration rate increased the employment growth of the low-skilled by 1.6% and high-skilled by 4.6%.

Given the predictions of the model, one would expect larger elasticities for the low-skilled. How to explain this? First, it is important to note that different skill groups have different degrees of wage rigidities and elasticities in labor supply responses. As discussed in Dustmann et al. (2017a), these

⁹⁰I assume that the relevant split in terms of production abilities is between college and non-college-educated workers. This is consistent with Goldin and Katz (2007); Card (2009); Ottaviano and Peri (2012); Docquier et al. (2014) who find high substitutability between workers with no schooling and high school degree, but small substitutability between those and workers with a college education.

⁹¹Estimate for the high-skilled is significant at 11%.

differences can create “perverse” effects in which the group that experiences the greatest shock may not be the group that benefits the most in terms of wages or employments. Second, these elasticities measure the growth rates in the employment and not the increase in the number of workers. For instance, given the average number of low-skilled workers (2196) and high-skilled (486), a 1% increase in the out-migration rate adds around 41 low-skilled and 23 high-skilled workers.

3.9.2 By Demographics

Out-flow of the immigrants can have a different impact depending on the demographic characteristics of the individual. Table 14 provides a more detailed analysis by investigating how the out-migration of the immigrant affects labor market outcomes of different demographic groups. The estimation is identical to the baseline specification, which links the *overall* out-migration of (male) immigrants to employment and wage growth of each group.

Table 14: Wages and Employment by Demographic Groups

Panel A: Change Wages					
	Native Male	Native Female	under 30	30-40	above 50
Net Outflow Rate	1.4104 (0.629)**	1.2763 (0.634)**	1.1551 (0.608)*	1.3706 (0.690)**	1.5596 (0.800)*
N	250	250	250	250	250
Adj R2	0.64	0.70	0.69	0.50	0.59
KP F-Stat	18.50	18.29	18.99	18.99	17.00
Panel B: Change Employment					
	Native Male	Native Female	under 30	30-40	above 50
Net Outflow Rate	2.5822 (0.786)***	3.3757 (1.046)***	6.1559 (2.298)***	0.9597 (0.566)*	2.3923 (0.708)***
N	250	250	250	250	250
Adj R2	0.37	-0.13	0.02	0.33	0.28
KP F-Stat	18.50	18.29	18.99	18.99	17.00

The wages and employment growth in the previous periods, standardized by initial employment. The table reports OLS and 2SLS estimates for the impact of net change in immigrant population in an area, measured as the decrease in the number of working-age male immigrant population between t and $t - 1$ as of total working-age male population in year $t - 1$, on native local wage and employment growth in the aggregate. The unit of observations are provinces. Regressions are weighted by group-specific employment in year $t - 1$ and include year fixed-effects. Standard errors are clustered at province-level. Data source: MCVL and INE.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel A presents the results for wage growth, while Panel B reports results for employment. In terms of wages, all groups benefit from the departure of the immigrants, although the estimated elasticities are not statistically different from one another. Given the wage rigidities in Spain, it is reasonable that there are no significant differences.

The outflows accelerate employment growth of all groups as well. While males (Column 1) and females (Column 2) benefit equally from outflows, there are significant differential effects by the age

groups. The elasticity for natives who are under the age of 30 (Column 3) is two to three times larger than other groups, and the difference is statistically significant. There are two possible explanations for this stronger effect. First, as predicted by the model, natives who have the highest substitutability with the immigrants should benefit the most from their departure. Given that the departing immigrants were young and more likely to be competing for jobs that are held by young natives (due to downgrading), this result is understandable. Second, young natives in Spain suffer from high unemployment rates. Thus increase in their employment should generate mechanically (i.e., smaller denominator) higher growth rates in group employment. In reverse, natives who are between age 30 to 40 (Column 4) benefit the least from the outflows, possibly due to their higher labor market integration.

3.9.3 EU versus non-EU outflows

As discussed in Appendix Section 3.13.7, the Recession did not generate any apparent difference in flow patterns between EU and non-EU nationals. Still, given the differences in the labor market conditions of the two groups, the outflows could have different marginal effects. In order to test it, I split the outflows between those who were born in an EU member country vs. those who are born elsewhere. I construct separate outflow rates where I account outflows of these groups separately. In order to deal with endogeneity concerns, I build the instrument specific to each group.

Table 15: EU vs. Non-EU outflows

	Wage		Employment	
	2SLS	2SLS	2SLS	2SLS
Net Outflow Rate	2.0286 (0.609)***		2.4122 (0.752)***	
EU Outflows		2.1829 (0.983)**		1.0936 (0.636)*
Non-EU Outflows		1.8875 (0.621)***		4.1372 (1.524)***
N	250	250	250	250
r2	0.26	0.24	0.49	0.45
Cragg-Donalds Stat	34.73	16.43	34.73	16.43

The wages and employment growth in the previous periods, standardized by initial employment. The table reports OLS and 2SLS estimates for the impact of net change in immigrant population in an area, measured as the decrease in the number of working-age male immigrant population (EU vs. non-EU) between t and $t - 1$ as of total working-age male population in year $t - 1$, on native local wage and employment growth. Regressions are weighted by group-specific employment in year $t - 1$ and include year fixed-effects. Standard errors are clustered at province-level. Data source: MCVL and INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15 present these results. In columns 1 and 3, I present the benchmark results for comparison. In columns 2 and 4, I split the outflows between EU vs. Non-EU, and use instruments built separately for

each group. As can be seen in column 2, the outflow of EU-born foreigners has a larger point estimate, suggesting a more positive effect on the wages. Given the standard errors, it is not possible to say that they are statistically different. Results in column 4, however, show that departure of the non-EU-born foreigners has a much stronger effect on native employment. These results suggest that natives and non-EU-born foreigners have higher substitutability, compared to natives and EU-born foreigners. A potential explanation for these results is the composition of the outflows. Although not reported, most of the EU outflows were Romanian while non-EU outflows were from Latin America. Given the linguistic advantage of the Latin Americans and their high degree of labor market integration, these results are not surprising.

3.10 Mechanism

This section explores the mechanisms that are driving the results observed in the previous sections. I start by exploring the impact of out-migration on the geographic mobility of natives. In the following part of the section, I decompose the changes in the employment and wage margins to shed light on the underlying mechanisms.

3.10.1 Geographic Mobility of Natives

Natives can respond to the departure of immigrants from a local labor market through geographical mobility. Changes in the mobility patterns can impact both the native population levels and the growth rate of the native workforce.

In addition to its importance as a mechanism, understanding the mobility patterns is also crucial for identification. When estimating the local average impact of immigration outflows, one needs to take into account that changes in population in a given area affects the whole regions system equilibrium. The relocation of the population across areas within a country would hinder the identification of any area-level effects, as the effects dissipate throughout the country. Similarly, if out-migration of the immigrant triggers outflows of natives in parallel, then the coefficient would be overestimating the effect of immigrants' departure. If on the other hand, it triggers an inflow of natives, the final effect might be underestimated or null. This means that the average effect of out-migration on local labor markets would be affected by native mobility.⁹²

I test the impact of out-migration on native mobility following the empirical strategy proposed by

⁹²The importance of native displacement is not limited to simply understanding the direct question of natives' adjustment margin. Through mobility of natives, the effects of immigration can be diffused to other labor markets, which would invalidate the results obtained from cross-regional analyses (Lewis and Peri, 2015). Thus it has been an important element in the long-standing debate on whether immigration reduces the employment opportunities and wages of natives. According to Borjas (2006), the native migration response attenuates the measured impact of immigration on wages in a local labor market by 40 to 60 percent. Failure to account for this mechanism is given as an explanation for lack of robust estimates on the impact of immigration on wages in the US. Some papers however find no or little displacement effect (Card, 2001; Card and DiNardo, 2000). Replicating various methodologies in the literature Peri and Sparber (2011) find evidence against the existence of native displacement due to immigration.

Peri and Sparber (2011). Specifically, I use a normalised change in native population as the dependent variable and use exactly the same right-hand-side elements that I use in my main analysis (12 or 13):

$$\Delta native_{j,t} = \lambda \Delta foreign_{j,t} + \alpha_t + \varepsilon_{j,t} \quad (18)$$

where $\Delta native_{t,j}$ is the net-change in native-born population between two periods, divided by the province initial population.⁹³ The sign and size of λ captures the relationship between immigrant outflows and natives' location choice. If the estimated λ is positive, this would indicate that native population is increasing in areas where immigrants are leaving. If natives leave areas along with immigrants then estimated λ will be negative. Table 16 shows the results of the estimation of equation 18. Columns 1-3 show the OLS and 4-6 show the 2SLS results, where I use the same instrument.

Table 16: Impact of out-migration on native mobility between 2009-2014

	OLS			IV		
	Male	Female	Total	Male	Female	Total
Net Outflow Rate	0.1703 (0.083)**	0.1709 (0.079)**	0.1654 (0.080)**	0.3831 (0.154)**	0.3763 (0.154)**	0.3788 (0.153)**
Bartik	Yes	Yes	Yes	Yes	Yes	Yes
N	250	250	250	250	250	250
Adj R2	0.26	0.24	0.25	0.19	0.18	0.19
KP F-Stat				15.75	16.75	15.76

The table reports OLS and 2SLS estimates for the impact of net change in immigrant population in an area, measured as the decrease in the number of working-age male immigrant population between t and $t - 1$ as of total working-age male population in year $t - 1$, on the mobility of pre-existing native working age population during the same period. I use the methodology from Peri and Sparber(2011). Regressions are estimated at the yearly level, across 50 provinces. Regressions are weighted by group-specific employment in the year $t - 1$ and include year fixed-effects. Standard errors are clustered at province-level. Data source: MCVL and INE.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

I find a positive and significant impact of out-migration flows on the native population, both in the OLS and in the IV results, both for males and females. These results show that the outflow of immigrants increases the net native population. These estimates predict that for every 3 immigrant that leave a given province, 1 *additional* native locates in the same province. This elasticity is exact the same as those found in Boustan et al. (2010); Dustmann et al. (2017a); Fernández-Huertas Moraga et al. (2017); Monras (2018).⁹⁴

Although informative, these results do not tell us anything about how the native population is increasing. Natives can respond by decreasing in-migration from other areas, increasing out-migration from the area, or both. In order to answer this question, I use the EVR data and calculate arrival and departure rates for the natives.

⁹³ $\Delta native_{j,t} = \frac{\Delta N^{native}_{j,t} - N^{native}_{j,t-1}}{N^{native}_{j,t-1} + N^{foreign}_{j,t-1}}$

⁹⁴When I repeat the exercise in 5-year long-differences (LD) from 2009 to 2014, I get similar results. Results can be provided upon request.

Specifically, I define the arrival rate to province j at time t as follows:

$$Arrival_{j,t} = \frac{Arrivals_{j,t}}{Native_{j,t}}$$

where $Arrivals_{j,t}$ denotes the number of natives that live in j at time t and were living somewhere else at time $t - 1$. Similarly, I define the departure rate from a province j as:

$$Departure_{j,t} = \frac{Departure_{j,t}}{Native_{j,t}}$$

where $Departure_{j,t}$ denotes the number of individuals that lived in j at time $t - 1$ and were living somewhere else at time t .

In Table 17, I show this decomposition at the province level. The arrival and exit rates can differ between provinces due to structural differences or province-specific trends. In order to net out these differences, I include province fixed-effects along with time fixed effects. Given that outflow rates are stronger in areas which were hit harder by a demand shock, I include the Bartik control.

Native populations in provinces grow faster in areas where the out-migration rate is higher, and it is due to an increase in the arrival rates.⁹⁵ This result echoes the findings of Dustmann et al. (2017a); Ortega and Verdugo (2016); Sanchis-Guarner (2017); Monras (2018) who report a decline in native inflows in response to immigration, while finding no effect on outflows; or Dao et al. (2017) and Monras (2018) who find that population adjustments during a local demand shock mainly come from changes in the arrival rates rather than departure rates.

Table 17: Impact of out-migration on native mobility between 2009-2014

	Arrival Rate	Departure Rate	Net Rate
Net Outflow Rate	0.1054 (0.042)**	-0.0258 (0.021)	0.1311 (0.043)***
Bartik	Yes	Yes	Yes
Prov FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	250	250	250
KP F-Stat	12.44	12.44	12.44

The table reports 2SLS estimates for the impact of net change in immigrant population in an area, measured as the decrease in the number of working-age male immigrant population between t and $t - 1$ as of total working-age male population in year $t - 1$, on the mobility rates of pre-existing native working age population during the same period. Regressions are estimated at the yearly level, across 50 provinces. Regressions are weighted by group-specific employment in the year $t - 1$. All regressions include year and province fixed-effects. Standard errors are clustered at province-level. Data source: INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

⁹⁵The EVR data distinguishes moves within two provinces and those from and to abroad. Here I use the total of both domestic and international moves for simplicity. However, the results are mainly driven by domestic moves.

These results show that out-migration of the immigrants attracted natives to these areas. This could be suggestive that natives and immigrant have a certain degree of substitutability. This also means that the estimated impact of the out-migration on the change in the average wages between 2009 and 2014 includes a downward bias due to the arrival of the natives, and thus presents a lower bound for the positive effects.

3.10.2 Margins of the Employment Effects

The positive employment effects presented so far are based on the average changes across local labor markets. These observed effects can be due to the increased entry of the natives into employment, lower probability of separation, or exit from the labor market.

The overall employment effects reported can be decomposed between movements from and to nonemployment (i.e., inactive or unemployed). Specifically, $\frac{L_{j,t}^{Native} - L_{j,t-1}^{Native}}{L_{j,t-1}^{Native}} = \frac{Entries_{j,t}^{Native}}{L_{j,t-1}^{Native}} - \frac{Exits_{j,t}^{Native}}{L_{j,t-1}^{Native}}$ where $Entries_{j,t}^{Native}$ is the number of natives employed in area j in year t but not in year $t - 1$, while $Exits_{j,t}^{Native}$ captures those natives who were employed in $t - 1$ but not in t . In order to test this, I first calculate the entry and exit rates for each skill group.⁹⁶ Then, I formally test it by using these rates as the dependent variable. These entry and exit rates can differ between provinces due to structural differences or province-specific trends. In order to net out these differences, I include province-fixed effects along with time-fixed effects. Hence the estimated parameters give the deviation from the province average.

Table 18 reports the results for entry (columns 1-3) and exit rates (columns 4-6) for different skill groups. These results indicate that an increase in the outflow rate accelerates the entry from nonemployment to employment for both skill groups. An increase of 1% in the outflow of the immigrant population accelerates the entry rate of low-skilled and high-skilled natives into employment by 2.4% and 2%, respectively. Results in columns 4-6 suggest that outflows are also decelerating the exit rates of low-skilled and high-skilled by 1.3% and 0.4%, although these coefficient are statistically insignificant.

These results show that, for all skill groups, most of the employment effects go through increased entries rather than decreased exits. This suggests that the outflow of immigrants benefit the “outsiders” who are more likely to enter to employment, rather than “insiders” who are already employed. This results mirror the effects found in Dustmann et al. (2017a) who show that inflow of Czech workers into German municipalities have reduced the entries of natives from nonemployment to employment, while those who were already employed were not affected.

⁹⁶I calculate the entry and exit rates, by dividing the number of newly employed workers normalised by the number of employed workers in the area.

Table 18: Entries versus Exits: Employment

	Entry			Exit		
	All	Low Skilled	High Skilled	All	Low Skilled	High Skilled
Net Outflow Rate	2.3223 (0.875)***	2.4515 (0.938)***	1.9569 (0.990)**	-1.2718 (1.063)	-1.3810 (1.178)	-0.3503 (1.208)
Prov FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	250	250	250	250	250	250
KP F-Stat	22.92	22.92	22.92	22.92	22.92	22.92

Dependent variables are entry rate (columns 1-3) and exit rate (columns 4-6). The unit of observations are provinces. Regressions are weighted by total employment in year $t - 1$, include province and year fixed-effects. Standard errors are clustered at province-level.

Data source: MCVL and INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.10.3 Margins of the Wage Effects

Average wage growth observed in the previous section can be driven by three different channels. The increase in the average wages can be due to higher entry wages, wage growth of stayers or the exit of the least productive workers. In order to analyse this, I use the panel dimension of my data and decompose the wage growth separately for those who enter employment (“*New Entries*”) and for those who were employed in both periods (“*Stayers*”).⁹⁷ Table 19 presents these results. column 1 to 3 show the change in the entry wages while column 4 to 6 show the results for those who remained employed in both periods.

Results in the first three columns indicate that the departure of the immigrant increased the wages of those who entered employment, although the results are statistically not significant. It is important to note that these insignificant results are driven by two opposing forces. As discussed in the model, the departure of immigrants decreases competition in the labor market, thus pushing up wages in the labor market. On the other hand, the departures can also potentially increase the entry of less productive workers into employment (negative selection), which could lower the average wages.

Results in columns 4 to 6, show that higher departure rates increased the average wages for those who were employed in both periods. Given that in this specification the composition of workers is fixed between two periods, these results are entirely driven by decreased competition in the labor market due to the departure of the immigrants.

⁹⁷Here I define a worker as a “*Stayer*” if that person was employed for two successive periods. I do not however distinguish whether the individual remained in the same job or changed.

Table 19: New Entries vs Already Employed: Wages

	New Entries			Stayers		
	All	Low Skilled	High Skilled	All	Low Skilled	High Skilled
Net Outflow Rate	1.8395 (1.202)	1.4190 (1.219)	1.7956 (4.161)	1.1957 (0.439)***	0.7391 (0.325)**	1.8437 (0.578)***
N	250	250	244	250	250	250
KP F-Stat	38.94	36.76	40.74	29.85	27.07	35.98

The wage and employment growth in the previous periods, standardized by levels in the base period. The explanatory variable is the share of immigrant in the total working age population. The unit of observations are provinces. Regressions are weighted by total employment in year $t - 1$, include province and year fixed-effects. Errors are clustered at province-level. Data source: MCVL and INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Daily Wage Rate versus Days Worked

Previous sections presented results on the changes in monthly earnings. These earnings are composed of daily rates that workers received per day of work, multiplied by total days worked in that month. In this section I test whether the positive earnings effects are due to an increase in daily wage rates, number of days worked, or both. Table 20 reports results for both margins. Columns 1-2 show that the outflow of the immigrant increased average daily wages for both skill groups. Columns 3-4 report the changes in the average number of days worked in a month. The results suggest that departure of immigrant increased average daily wages but not the number of days worked for natives.⁹⁸ These results mirror the findings in Edo (2016) who find exactly the effects in exactly the same margins but in the opposite direction due to increased immigration in the French labor market.

Table 20: Daily Wages vs Days Worked in a month

	Daily Wage		Days Worked	
	Low Skilled	High Skilled	Low Skilled	High Skilled
Net Outflow Rate	1.1112 (0.332)***	1.5213 (0.844)*	0.1521 (0.139)	0.0697 (0.260)
N	250	250	250	250
KP F-Stat	16.74	17.90	16.74	17.90

The unit of observations are provinces. Regressions are weighted by total employment in year $t - 1$ and include year fixed effects. Standard errors are clustered at province-level.

Data source: MCVL and INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

⁹⁸These results are not driven by compositional changes. Inclusion of changes in average schooling and age of the employed does not change results. Results can be provided upon request.

3.10.4 Margins by Contract Types

As shown in the previous sections, the departure of the immigrants have increased both the employment and the wages of workers of both skill groups, which confirms the standard economic theory and the results in Equation 13. However, as I discuss in the Section 3.2.2, the prediction of the model and the margins of adjustment depend on the degree of wage rigidity.

Spain has a dual labor market which creates important differences between indefinite and fixed-duration contracts in terms of firing costs and wage setting (Bentolila et al., 2012a,b).⁹⁹ As shown in De la Roca (2014) while indefinite contracts provide higher protection they also generate greater wage rigidities and thus lower cyclicalities. Fixed-term contracts, on the other hand, provide higher flexibility in terms of wages. Since the duration of fixed-term contracts is relatively short, and their termination has no cost, firms can adjust their demand through these contracts. As the wages of these contracts are less influenced by the institutional framework, the wage levels better reflect the cyclicalities.¹⁰⁰ Given these differences in the degree of wage rigidities, the type of job contracts may affect the responsiveness of wages and employment to outflow of immigrants.

In order to test how outflows of immigrants impact the wage and employment growth of natives, I decompose the workers into two groups by type of job contract: those who have a fixed-duration contract and those who have an indefinite (permanent) contract.¹⁰¹ I repeat the regression for both groups separately and present the 2SLS results in Panel A of Table 21. For each contract type, I present wage growth (columns 1 and 2) and employment growth (columns 3 and 4). In order to prevent the results from being driven by differences in labor demand or changes in the composition of the workers, I include Bartik and demographic controls.

⁹⁹Workers under permanent contracts benefit from a high level of employment protection through generous severance payments and legal defense in case of a firing event. Workers under temporary contracts have much lower severance payments and do not face legal proceedings when the contracts expires. As a result, workers in permanent contracts enjoy high protection and bargaining power, while workers in temporary contracts earn lower wages and suffer from high turnover rates and low levels of job tenure.

¹⁰⁰For instance, in Spain De la Roca (2014) shows that wage cyclicalities for workers under temporary contracts is twice as large for workers under permanent contracts.

¹⁰¹Indefinite contracts include the *contrato indefinido ordinario* and the *contrato de fomento de al contraction indefinida*. Temporary job contracts are *emploi interimaire* or *contrato temporal*.

Table 21: Wages and Employment: Contract Types

Panel A: Average Wages and Employment				
	Wages		Employment	
	Indefinite	Fixed	Indefinite	Fixed
Net Outflow Rate	-0.2312 (0.397)	4.1571 (1.348)***	3.8046 (1.380)***	-1.0169 (1.517)
Bartik	Yes	Yes	Yes	Yes
Demographic	Yes	Yes	Yes	Yes
N	250	250	250	250
KP F-Stat	11.27	19.51	11.27	19.51
Panel B: Wages by Cohorts				
	New Entries		Already Employed	
	Indefinite	Fixed	Indefinite	Fixed
Net Outflow Rate	-4.0547 (2.739)	2.7734 (1.853)	0.6037 (0.288)**	3.8769 (1.797)**
Bartik	Yes	Yes	Yes	Yes
Demographic	Yes	Yes	No	No
N	250	250	250	250
KP F-Stat	15.12	22.50	23.30	15.30

The table reports 2SLS estimates for the impact of net change in immigrant population in an area, measured as the decrease in the number of working-age male immigrant population between t and $t - 1$ as of total working-age male population in year $t - 1$, on the wage and employment growth by type of contract. Regressions are estimated at the yearly level, across 50 provinces. Regressions are weighted by group-specific employment in year $t - 1$ and include year fixed-effects. Standard errors are clustered at province-level.

Data source: MCVL and INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns 1 and 2 show that the outflow of immigrants increased the wages of workers under fixed-term contract positively and significantly while not affecting the wages of those with indefinite contracts. Columns 3 and 4 show that provinces where outflows were higher, the number of workers with indefinite contracts increased, while the number of fixed contracts was unaffected. These margins of adjustment echo the findings of Edo (2016) for France.

These results are however based on the average effect of out-migration on both newly recruited workers and those who were in the labor market. In order to understand the underlying mechanism, in Panel B of Table 21 I further decompose the wage growth of the worker by those who entered the labor market vs. those who were already in the labor market. The results are informative. Column 1 and 2 show that the out-migration did not affect the wages of the newly recruited individuals, which is consistent with the previous findings. On the other hand, as can be seen in columns 3 and 4, those who were employed in the previous period saw their wages increase due to the departure of the immigrants. The differences in the estimated coefficients between column 3 and 4 show that increases are mainly driven by the fixed-contracts. These elasticities are in line with De la Roca (2014), who finds that wage cyclicality in Spain for temporary contracts is twice as large as for workers under indefinite contracts.

These results complement the findings in the section. While “outsiders” benefited from the employ-

ment margin through increased entries to employment, “insiders” benefited from wage growth.

3.11 Understanding the Magnitude of the Estimates

The estimated elasticities of the wage and employment effects may seem large. Results in Table 3 mean that 1% decrease in the working age male population due to net outflow of the immigrant population increases the wages of the employed natives by 2% and the employment by 2.4%. Given the averages presented in Table 1, for every 3160 immigrant who leave the province, natives’ monthly wages increase by 32 euros and 64 natives find full-time formal employment in the same year. Of these 64 jobs, 23 of them are high-skilled, while 41 of them are low-skilled. In addition to the full-time employment, outflows also increase part-time employment (See Appendix Section 3.13.17 for the regression results). More precisely, departure of the immigrants also add 15 part-time workers, of which 3 of them are high-skilled and 12 of them are low skilled.

It is important to note that these estimates are obtained in a context of economic recession. During my period of analysis, on average 43 low-skilled and 4 high-skilled jobs in the formal sector were destroyed, and average wages decreased by 2% annually in each province. Given the overall decline in employment and wages, the outflows from provinces slowed down the drop in native employment as predicted in the model. In other words, the mobility of the immigrant provided a cushioning effect to the natives.

These estimations are, however, valid for formal sector employment and wages. Whether similar effects can be expected in the informal sector is another question which is beyond the scope of this paper. Moreover, given the strict sample selection criteria (e.g., 30 days of employment within the month, minimum 30 days of employment within the year) it is likely that the employment numbers present the lower bound of the positive effects.

These results should be considered within its specific context. Spain is a country where the immigrant population has an average substitutability with native population, which is much higher than many countries. For instance, immigrants who left were from Spanish speaking countries or other European countries, thus were close substitutes to the natives at least in terms of linguistics capabilities or skills. High substitutability between immigrants who left and natives who stayed potentially can explain part of the positive effects. That is why, although the conclusions derived from this case study can be generalized to other migration contexts, the effects could be of smaller magnitudes when immigrants and natives of similar observable skills (i.e., education and age) are imperfect substitutes or complement in production.

The positive wage effects are in line with those found by Borjas and Aydemir (2007), Mishra (2007), Elsner (2013) and Dustmann et al. (2015), although they use different empirical strategies, study different countries and focus on labor market outcomes of natives. I further find that the departure of the immigrants increase the native population in the area, through increased inflows, which mirrors the effect found in the literature on the displacement effect of immigration (e.g. Amior, 2017). My results

show that outflow of the immigrant population increases both recruitment of natives who were previously not employed. This observation indicates that “outsiders” are the group that benefits from the departure of the immigrant the most both in terms of wages and employment. This finding echoes the finding of Dustmann et al. (2017a), who find that the immigration of the immigrant workers impacts only “outsiders” while “insiders” are not affected potentially due to the labor market institutions.

3.12 Concluding Remarks

This paper documents the impact of outflow of immigrants on the wages and employment of native workers in Spain during the Great Recession. Using administrative data and municipal population registers, I find that the out-migration accelerated the wage and employment growth for natives of all skill and demographic groups. I find that the departure of immigrants increased geographic inflows of natives from other areas, increased the entry to employment from nonemployment, and improved the wages of those who were already employed.

Given the context of economic contraction, these findings show that through their higher mobility, immigrants diffuse the incidence of local shocks and cushion the natives during an adverse shock. In a context of wage stagnation and employment losses, the departure of immigrants, improved the wage growth while dampening the drop in the employment. As the locations which were hit harder by the crisis also lost more immigrants, this smoothing effect contributes as a mechanism for equilibrating differences across labor markets. This suggests that migrants do not only grease the wheels during the good times (Borjas (2001)), but also during bad times by leaving locations that are hit by a negative shock. This finding underlines an important benefit of immigration that has been explored little in the literature. This is particularly important given the concerns about the relative lack of mobility of natives and especially among less-skilled workers as it leads to significant divergence in local unemployment rates and workers’ earnings across local labor markets (Bound and Holzer, 2002; Cadena and Kovak, 2016; Dao et al., 2017).

These findings also shed light on how rigid labor markets adjust to changes in labor supply. It shows that labor market rigidities determine whether the effects operate through the employment or wage margin. This suggests that my findings have implications beyond the immigration literature and contribute to the understanding of how labor markets respond to local shocks.

This paper provides evidence on the impact of immigrant mobility on the local labor market outcomes of natives. It does not, however, focus on the individual heterogeneity of either of the groups and address the individual selection component. Better understanding of the selection due to individual ability both for departing immigrants as well as natives is yet to be addressed and remains an important question for future research.

3.13 Appendix

3.13.1 Derivation of the Firm's Demand Curve

In perfect competition, factors of production are paid according to their marginal productivity. In such set-up, firms maximize their profits for given output and input prices.

$$\pi = A[\theta_U L_U^\beta + \theta_S L_S^\beta]^{\frac{1}{\beta}} - w_S L_S - w_U L_U$$

First-order condition:

$$\frac{\partial \pi}{\partial L_g} = A \frac{1}{\beta} \beta L_g^{\beta-1} \theta_g L^{1-\beta} - w_g = 0$$

$$w_g = A(\theta_g) L_g^{\beta-1} L^{1-\beta}$$

This relationship is independent of the sign of β . The magnitude of the effect, however, depends on the value of β (see graph with parameters). I take the logarithm of the demand function for one group (as both give the same result):

$$\log w_g = \log(A) + \log(\theta_g) + (\beta - 1)(\log L_g) + (1 - \beta)(\log L)$$

Calculating the total differential with respect to variations of employment of workers within the skill group (i.e. L_S) and of the total employment (i.e. L) gives me this:

$$d \log w_g = (\beta - 1) d \log L_g + (1 - \beta) d \log L + d \log A \quad (19)$$

I take the total differential of L :

$$L = [\theta_U L_U^\beta + \theta_S L_S^\beta]^{\frac{1}{\beta}}$$

$$d \log L = \frac{1}{L^\beta} \frac{1}{\beta} [d L_U \theta_U \beta L_U^{\beta-1} + d L_S \theta_S \beta L_S^{\beta-1}]$$

$$d \log L = \frac{1}{L^\beta} [d L_U \theta_U L_U^\beta L_U^{-1} + d L_S (\theta_S) L_S^\beta L_S^{-1}]$$

$$d \log L = \frac{1}{L^\beta} \left[\frac{d L_U}{L_U} \theta_U L_U^\beta + \frac{d L_S}{L_S} (\theta_S) L_S^\beta \right]$$

$$d\log L = \frac{1}{L^\beta} [d\log L_U \theta_U L_U^\beta + d\log L_S (\theta_S) L_S^\beta]$$

$$d\log L = \frac{d\log L_U \theta_U L_U^\beta}{L^\beta} + \frac{d\log L_S (\theta_S) L_S^\beta}{L^\beta}$$

$$d\log L = S_U d\log L_U + S_S d\log L_S \quad (20)$$

where $S_U = \frac{\theta_U L_U^\beta}{L^\beta}$, $S_S = \frac{\theta_S L_S^\beta}{L^\beta}$

Totally differentiating L_g :

$$L_g = L_g^N + L_g^M$$

$$dL_g = dL_g^N + dL_g^M$$

$$\frac{dL_g}{L_g} = \frac{dL_g^N}{L_g^N} + \frac{dL_g^M}{L_g^M}$$

$$\frac{dL_g}{L_g} = \frac{dL_g^N}{L_g^N} \frac{L_g^N}{L_g} + \frac{dL_g^M}{L_g^M} \frac{L_g^M}{L_g}$$

$$d\log L_g = \theta_g^N d\log L_g^N + \theta_g^M d\log L_g^M \quad (21)$$

where $\theta_g^N = \frac{L_g^N}{L_g}$, $\theta_g^M = \frac{L_g^M}{L_g}$

Assuming that I am interested in the outcomes of skilled workers (i.e. $g = s$). I plug Equation 20 into Equation 19 for $d\log L$:

$$d\log w_S = (\beta - 1)d\log L_S + (1 - \beta)(S_U d\log L_U + S_S d\log L_S) + d\log A$$

$$d\log w_S = [(\beta - 1 + (1 - \beta)S_S)]d\log L_S + (1 - \beta)(S_U)(d\log L_U) + d\log A$$

Plug in Equation 21 for $d\log L_g$:

$$d\log w_S = [(1 - \beta)(S_s - 1)][d\log L_S^N \theta_S^N + d\log L_S^M \theta_S^M] + (1 - \beta)(S_U)[d\log L_U^N \theta_U^N + d\log L_U^M \theta_U^M] + d\log A$$

Isolating $d\log L_S^N$ gives me the labor demand (Equation 4) in the paper.

3.13.2 Derivation of Labor Supply

Starting from:

$$L_g = N_g f_g(w_g, w'_g) + L_g^M$$

$$L_g^N = N_g f_g(w_g, w'_g)$$

$$\log(L_g^N) = \log(N_g f_g(w_g, w'_g))$$

$$d\log(L_g^N) = \frac{\partial(N_g f_g(w_g, w'_g))}{(N_g f_g(w_g, w'_g))} \frac{dw_g}{w_g} \frac{w_g}{w_g}$$

$$d\log(L_g^N) = d\log w_g \frac{w_g}{\partial w_g} \frac{\partial(N_g f_g(w_g, w'_g))}{(N_g f_g(w_g, w'_g))} \quad (22)$$

This gives me the labor supply in Equation 6 in the text.

3.13.3 Derivation of Equilibrium Wage and Employment Responses under Flexible Wages

The equilibrium wage and employment responses are determined by the two skill-specific labor demand curves:

$$d\log w_S = (\beta - 1)d\log L_S + (1 - \beta)d\log L + d\log A \quad (23)$$

$$d\log w_U = (\beta - 1)d\log L_U + (1 - \beta)d\log L + d\log A \quad (24)$$

and two skill-specific supply curves:

$$d\log L_S^N = \eta_S d\log w_S \quad (25)$$

$$d\log L_U^N = \eta_U d\log w_U \quad (26)$$

where $d\log L$ is given by Equation 20. By plugging 25 and 26 in 23 and 24, I obtain:

$$dlogw_S = [(1-\beta)(S_s-1)][\theta_S^N dlogL_S^N + \theta_S^M dlogL_S^M] + (1-\beta)(S_U)[\theta_U^N dlogL_U^N + \theta_U^M dlogL_U^M] + dlog(A) \quad (27)$$

$$dlogw_U = [(1-\beta)(S_u-1)][\theta_U^N dlogL_U^N + \theta_U^M dlogL_U^M] + (1-\beta)(S_S)[\theta_S^N dlogL_S^N + \theta_S^M dlogL_S^M] + dlog(A) \quad (28)$$

Regroup and solve 27 for $dlogw_S$:

$$dlogw_S = \frac{(1-\beta)(S_U)[\eta_U(dlogw_u)\theta_u^N + dlogL_U^M\theta_U^M] + [(1-\beta)(S_s-1)]dlogL_S^M\theta_S^M + dlogA}{1 - [(1-\beta)(S_s-1)\eta_s\theta_S^N]} \quad (29)$$

$$dlogw_U = \frac{(1-\beta)(S_S)[\eta_S(dlogw_S)\theta_S^N + dlogL_S^M\theta_S^M] + [(1-\beta)(S_u-1)]dlogL_U^M\theta_U^M + dlogA}{1 - [(1-\beta)(S_u-1)\eta_u\theta_U^N]} \quad (30)$$

Plugging 30 into 29 and placing all terms over a common denominator then yields:

$$\begin{aligned} dlogw_S = & \frac{(1-\beta)(S_U)[(\eta_U\theta_U^N)(1-\beta)(S_S)(\eta_S\theta_S^N)(dlogw_S) + (\eta_U\theta_U^N)(1-\beta)(S_S)(\theta_S^M)(dlogL_U^M) + \theta_U^M dlogL_U^M]}{(1 - [(1-\beta)(S_S-1)]\eta_s\theta_S^N)((1 - ((1-\beta)(S_U-1)]\eta_U\theta_U^N))} \\ & + \frac{[[(1-\beta)(S_s-1)](dlogL_S^M)\theta_S^M](1 - ([\beta-1 + S_U - S_U\beta]\eta_U\theta_U^N))}{(1 - [(1-\beta)(S_S-1)]\eta_s\theta_S^N)((1 - ((1-\beta)(S_U-1)]\eta_U\theta_U^N))} \\ & + \frac{(dlogA)(1 - ((1-\beta)(S_U-1)]\eta_U\theta_U^N)}{(1 - [(1-\beta)(S_S-1)]\eta_s\theta_S^N)((1 - ((1-\beta)(S_U-1)]\eta_U\theta_U^N))} \end{aligned}$$

Solving for $dlogw_S$ gives:

$$dlogw_S^* = \frac{[((1-\beta)(S_S-1)) + (\eta_U\theta_U^N)[(1-\beta)^2(S_U)(S_S) - ((1-\beta)(S_U-1))(1-\beta)(S_U-1)]]\theta_S^M}{(1 - ((1-\beta)(S_S-1)]\eta_s\theta_S^N)(1 - ((1-\beta)(S_U-1)]\eta_U\theta_U^N)) - ((1-\beta)^2 S_U (\eta_U\theta_U^N)(S_S)(\eta_S\theta_S^N))} dlogL_S^M$$

$$+ \frac{[(1 - \beta)(S_U)(\theta_U^M)]}{(1 - ((1 - \beta)(S_S - 1)\eta_S\theta_S^N))(1 - ((1 - \beta)(S_U - 1)\eta_U\theta_U^N)) - ((1 - \beta)^2 S_U(\eta_U\theta_U^N)(S_S)(\eta_S\theta_S^N))} d\log L_U^M$$

$$+ \frac{(1 - (1 - \beta)(S_U - 1)\eta_U\theta_U^N)}{(1 - ((1 - \beta)(S_S - 1)\eta_S\theta_S^N))(1 - ((1 - \beta)(S_U - 1)\eta_U\theta_U^N)) - ((1 - \beta)^2 S_U(\eta_U\theta_U^N)(S_S)(\eta_S\theta_S^N))} d\log A$$

Plugging this into 25 and simplifying gives the equilibrium employment in Equation 11 in the paper.

3.13.4 Firms Adjustment During the Crisis

Wage Dynamics Network (WDN) survey¹⁰², conducted by the European Central Bank across the Eurozone, offers important insights into the effects of the crisis on firms and their adjustment during the 2010-2013 period.

According to WDN, Spanish firms reported that lack of demand was the main problem they had to deal with during the period.¹⁰³ Most of these firms perceived the shocks as partially persistent or permanent. Firms in service sectors observed stronger shocks compared to manufacturing and utilities, which is consistent with the higher decrease in domestic demand over this period. Finally, regardless of the sectors of activity or other firm characteristics (e.g., ownership), smaller firms reported the higher negative impact of the crisis as they suffered from both higher falls in demand and more significant difficulties in access to credit.

Despite the severe economic conditions, a significant fraction of Spanish firms did not reduce their labor costs between 2010 and 2013. While base wages increased moderately during this period, flexible wage component provided some room for a downward adjustment. However, as the weight of these components in total labor costs was rather small (4.2%), the scope for adjusting labor costs by these means was limited. Thus, very few firms, adjusted base wages or hours or hours per employee even when faced with strong fall in demand (Izquierdo and Jimeno (2015)). According to the survey only 1.6% in 2010, and 3.3% in 2013 reported having cut wages to their employees.¹⁰⁴

In contrast with other European countries, adjustments in working hours did not seem to play an important role (only 16.7% of firms reported a reduction in working hours over this period. Given the limited adjustments in wages and hours, the main instrument to accommodate the adverse shocks was a reduction in the number of employees. These reductions affected mostly employees under temporary contracts. Firms reduced their temporary and permanent employment, with the latter being the primary tool for especially those facing a significant drop in demand.

Reductions in employment were mainly through individual dismissals and non-renewal of temporary contracts, although the decrease in new hires was also quite frequent. The most recovery in new hires observed since the beginning of 2014 was mostly concentrated in temporary and part-time jobs. Most Spanish firms (close to 80%) report that uncertainty about economic conditions is an obstacle to permanent hires, although high labor costs (in the form of wages, non-wage labor costs) are also mentioned as a constraint for more than 60% of firms.

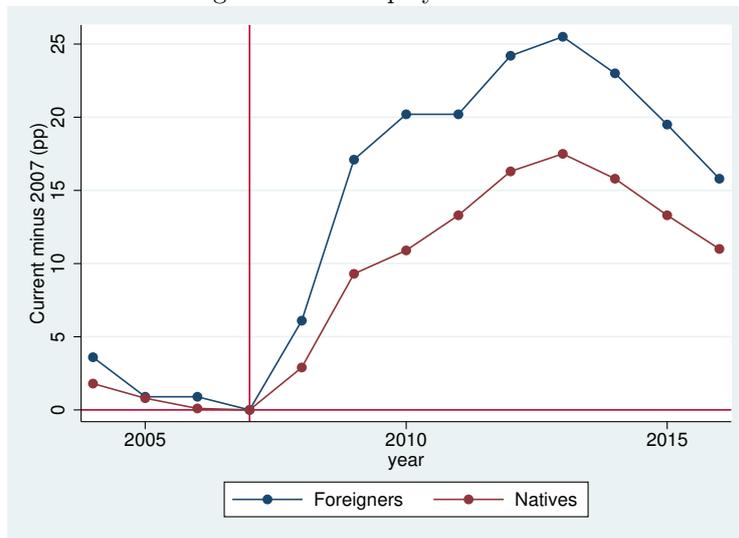
¹⁰²See Wage Dynamics in Europe: Final Report of the Wage Dynamics Network, available at http://www.ecb.europa.eu/home/pdf/wdn_finalreport_dec2009.pdf. The survey was conducted between September-December 2014, with a representative sample of 3,049 firms with more than 5 employees in the manufacturing, energy and market services sectors. As in previous WDN surveys in Spain, agriculture, construction and non-market services sectors were excluded. The questionnaire was about Spanish firms' perceptions about the size and the type of shocks affecting them during 2010-2013, and to what extent those shocks led to employment and wage changes.

¹⁰³According to the WDN more than 70% of firms said they underwent a strong or moderate decrease in demand, while only 15.4% of firms reported their demand to increase over this period. This fall in demand had mostly a national component since 65% of firms reported a decrease in domestic demand, and only 30.2% a decrease in external demand.

¹⁰⁴According to the previous wave of the WDN, wage cuts were practically inexistent in the pre-crisis period.

3.13.5 Differences in Unemployment Rates

Figure 15: Unemployment Rates



Data source: INE

3.13.6 Reconciling with Cadena and Kovak (2016)

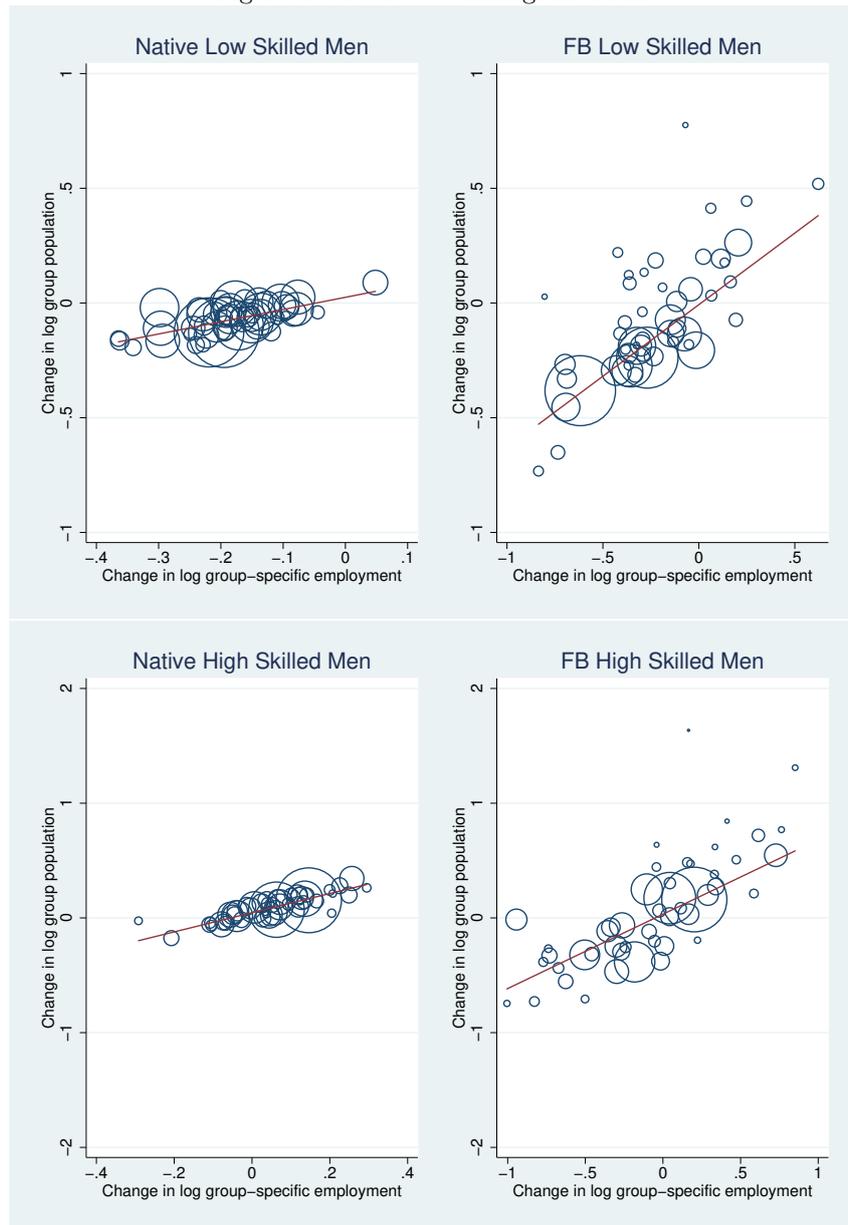
In recent work, focusing on the context the Great Recession in the USA Cadena and Kovak (2016) showed that immigrant workers (or at least low skilled migrants from Mexico) respond to changes in local labor demand through mobility across areas much more than their native counterparts. In consequence, the higher mobility of the immigrants reduces the spatial differences in employment and equilibrates the labor markets.

In this section, I replicate their key findings for 50 Spanish provinces for the period 2009 and 2014. I do this exercise to bridge the results of Cadena and Kovak (2016) with mine.

Population responses to employment shocks

I start by exploring the differences in the population responses to employment shocks. In Figure 11 of Section 3.3.3, I compare the differences between native-born and immigrant working-age men, for all skill groups. Figure 16 compares the mobility response of the low-skilled men (top panel), and of the high-skilled men (lower panel). These figures confirm that immigrant in Spain also respond much strongly to local demand shocks than the native-born population. Similar to the Cadena and Kovak (2016), these differences are especially striking for the low-skilled population.

Figure 16: Native vs immigrant Men



Changes are calculated as the long difference in logs from 2009 to 2014. Each circle corresponds to one of the 50 provinces. Observations are weighted by the group population in 2009. Data source: Labor Force Survey (EPA), INE.

Immigrant Mobility and Native Employment Outcomes

As shown previously, immigrant workers leave areas experiencing labor demand shocks at a much faster rate than the natives. Cadena and Kovak (2016) argue that this higher mobility smooths the employment effects of local labor demand shocks on the native population living in the area.

To test this mechanism, they study the relationship between local change in the employment rate and the local demand shock. They argue that the elasticity of employment rate to the local demand shock should be weaker in areas where the mobility is higher. More specifically, they measure this smoothing

effect by splitting the cities into two groups: those with above-median Mexican-to-population share before the crisis vs. those who have below-median Mexican-to-population share.

I repeat this exercise by splitting the provinces by their foreign-population share in 2007 and regress the following equation separately for each group:

$$\Delta \ln(\text{Emp.Rate}_{j,t}^g) = \beta^g \Delta \ln(\text{Employment}_{j,t}^g) + \delta_t + \varepsilon_{j,t}^g$$

where the dependent variable is the change in log of employment rate (employment to population ratio) for the group g (high-skilled, low-skilled etc.), located in province j , in time t . The independent variable is the change in the group-specific log employment and δ_t is the time fixed effects.

Table 22: Foreign-Mobility Smooths Employment Outcomes

	Above Median			Below Median		
	Native	Native HS	Native LS	Native	Native HS	Native LS
Ch. Native	0.4646 (0.060)***			0.7707 (0.063)***		
Ch. HS Native		0.2841 (0.039)***			0.3770 (0.047)***	
Ch. LS Native			0.3441 (0.049)***			0.6278 (0.053)***
Constant	-0.0047 (0.001)***	-0.0137 (0.003)***	-0.0066 (0.002)***	0.0005 (0.003)	-0.0099 (0.003)***	-0.0031 (0.002)
N	125	125	125	124	124	124
Adj R2	0.79	0.51	0.71	0.81	0.44	0.76

The table reports OLS estimates where the dependent variable is the change in log employment rates and independent variable is the change in the log employment of the relevant group. columns 1-3 present the results for provinces that have above-median immigrant population and columns 4-6 that have below-median immigrant population in 2007. The unit of observations are provinces. Regressions are weighted by group-specific population in year -1 and include year-fixed effects. Standard errors are clustered at province-level.

Data source: Labor Force Survey (EPA) and INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 22 reports the OLS results which confirm that the findings of Cadena and Kovak (2016) are also valid in Spain. The relationship between the employment rate and the labor demand shocks are much weaker in areas with high concentrations of immigrant workers. columns 3 and 6 show that the relationship is almost 50 percent weaker for low-skilled native workers, which is exactly the same rate found in Cadena and Kovak (2016) for the US cities. columns 2 and 5 show that the high-skilled also benefit from this mobility although, given the standard errors, it is not possible to reject the hypotheses that the elasticities are statistically different from each other.

Immigrant Outflows and Native Employment Outcomes

Cadena and Kovak (2016) establish two important facts: i) that migrants are more mobile than natives, ii) natives located in areas with the higher migrant population before the crisis experienced reduced job losses due to labor demand shock. In this paper, I bring these two mechanisms together and show

precisely how the migrant outflows can reduce the incidence of a negative demand shock on natives. In order to bridge my approach with that of Cadena and Kovak (2016), I repeat the exercise in Section 3.13.6, but split the provinces by the intensity of the outflows (treatment). Precisely, I measure the total outflows (normalized by the population in the base period) each province experienced between 2009 and 2014 and split the provinces into two groups according to the intensity of the outflows: those that experienced outflows that are above-median treatment vs. those with below-median treatment intensity.

Table 23: immigrant Outflows Smooths Employment Outcomes

	Above Median			Below Median		
	Native	Native HS	Native LS	Native	Native HS	Native LS
Ch. Native	0.4670 (0.064)***			0.7300 (0.057)***		
Ch. HS Native		0.3067 (0.043)***			0.3785 (0.046)***	
Ch. LS Native			0.3321 (0.059)***			0.5604 (0.061)***
Constant	-0.0037 (0.001)**	-0.0101 (0.002)***	-0.0067 (0.003)**	-0.0024 (0.003)	-0.0166 (0.004)***	-0.0065 (0.003)**
N	125	125	125	124	124	124
Adj R2	0.77	0.51	0.68	0.82	0.44	0.77

The table reports OLS estimates where the dependent variable is the change in log employment rates and independent variable is the change in the log employment of the relevant group. columns 1 to 3 present the results for provinces that have above-median immigrant outflow rates between 2009-2014 and columns 4 to 6 that have below-median immigrant outflow rates. The unit of observations are provinces. Regressions are weighted by group-specific population in year $t - 1$ and include year fixed-effects. Standard errors are clustered at province-level.

Data source: Labor Force Survey (EPA) and INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 23 presents the regression results. Overall the relationship between the employment rate and the labor demand shocks are much weaker in areas with higher intensity of immigrant outflows. The difference is especially significant for natives who are low-skilled. This findings confirm that my measure of outflows capture well the mechanism proposed in Cadena and Kovak (2016). I explore this mechanism in a more rigorous and causal framework in the rest of the paper.

3.13.7 Immigrant Population

EU vs. extra-EU Populations

Spain is part of the European Economic Area (which includes the EU) which provides freedom of movement across the area for the citizens of the member countries. The freedom of mobility can have significant consequences in the mobility choice of the individual. Following a contraction in the labor demand, a worker from a member country can leave Spain to another member state, with the knowledge of possible return in the future. On the other hand, a worker from a non-member country does not have freedom. First, he does not have the freedom to settle in another member country as his living permit is conditional on him finding a job before. Second, a worker who chooses to return home or to a third country has no guarantee of returning to Spain when the economy recovers.

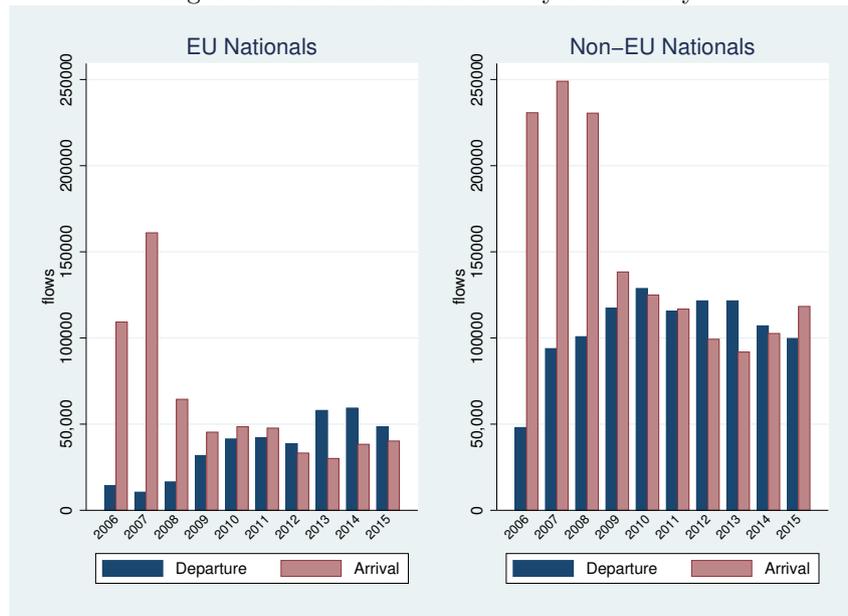
Given these differences, it is possible that two migrant workers, one from the EU and other from a non-EU country, may choose different mobility strategies as an adjustment to the crisis. In this section, I check whether the differences in mobility options have an impact on the mobility choice of the migrants and see whether migrants drive any mobility options from either one of the groups. To do so, I use the EVR and exploit its individual level data to distinguish the immigrants by their nationality, and split them into two by whether they have the nationality of an EU member state or not.

International arrivals and departures

International flows from/to Spain can be impacted differently depending on whether the individual has a EU passport or not. If the outflows are dominated by the mobility of those with EU passport, then the cushioning effects found in the main section are related to free mobility of workers within a customs union.

To investigate whether this is the case, I split the flows by nationality, between those with a passport belonging to one of the EU countries vs. those with a passport from extra-EU countries. Figures show the annual volumes of international arrivals and departures for immigrant working-age males. Figures do not display any striking difference between the two groups.

Figure 17: International Flows by Nationality



I can formally test that there is no change in the trend in immigrant stocks during this period from all countries of origin using Padrón. More specifically I use the following equation:

$$\ln(Foreign_{n,t}) = \delta_n + \delta_t + \beta EU_{n,t} + \varepsilon_{n,t}$$

where $Foreign_{n,t}$ is the total number of working-age immigrants from country n at year t . EU is a dummy variable that takes value equal to 1 if the country is member of the EU and thus its citizens benefit from free mobility at time t . If $\beta = 0$, it would be evidence that the mobility patterns are not different between EU and non-EU nationals. If instead $\beta \neq 0$ then it would mean that the mobility of the EU nationals changed during the period. I also include country fixed effects (δ_n) and time fixed effects (δ_t).

Table 24: Migrant Stocks EU vs. non-EU

	Ln(Pop)		Change in Ln(Pop)	
	2007-2014	2007-2016	2009-2014	2009-2016
EU27=1 × after=1	0.0270 (0.035)	0.0022 (0.042)		
EU27=1			-0.0489 (0.045)	-0.1053 (0.064)
N	903	1129	113	113
Adj R2	0.996	0.995	0.007	0.018

The table reports OLS estimates for the change in immigrant stocks in Spain, from country of origin n at year t . Regressions are estimated for 113 countries. columns 1 and 2 include time and country fixed effects, while columns 3 and 4 include only year fixed-effects. Standard errors are clustered at the country of origin level. Data source: INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 24 shows the results. column 1 shows that there is no systematic change in the stock of immigrants from EU member countries during the crisis. In this first column, I focus on the years between 2007, the year before the crisis and 2014, the final year of the Recession. In column 2, I change the end period to 2016, the last year where the net out-migration is positive. Regardless of the period, the results show that there is no change in the migration patterns during the Recession. In columns 3 and 4, I estimate the same regression in first differences, including the time fixed effect. Here the estimates are negative yet not statistically significant.

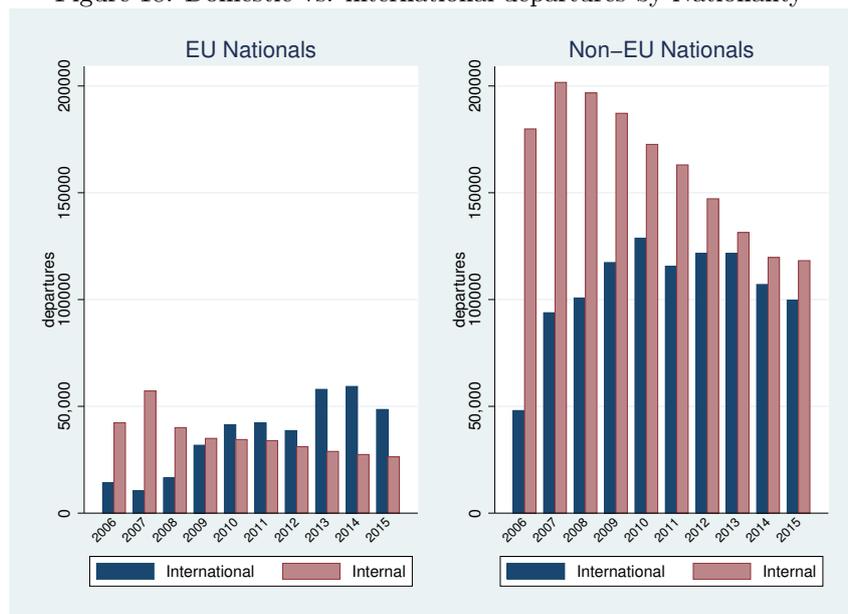
Both the figure and the results suggest that the mobility of the EU nationals did not differ from the mobility of the nationals of non-EU countries. This finding is also consistent with Basso et al. (2018), who do not find any difference in population elasticity to employment between EU and non-EU borns across EU countries during the period 2007-2016.

Domestic vs. International Departures

Similarly, one could wonder whether mobility choice between EU vs. non-EU nationals can be different in terms of international vs. domestic moves. For instance, an EU national can leave the economically distressed location and choose to move to another member State, while a non-EU national can decide to move domestically.

Using the EVR, I show the annual volumes of international and domestic departures for immigrant working-age males. In the left panel of the figure, it can be seen that the crisis has decreased the number of moves within Spain for the EU nationals marginally. The international departures however increased gradually, with a jump in 2013. Right panel, shows the moves by immigrant without the EU passport. The figure shows that, while international moves increased starting from 2008, the number of internal moves decreased dramatically. This reduction can be due to migrants choosing to reduce their mobility due to tighter labor market conditions as discussed in 3.3 (2017). The figures show that for both groups, departures with a destination abroad or within Spain changed similarly.

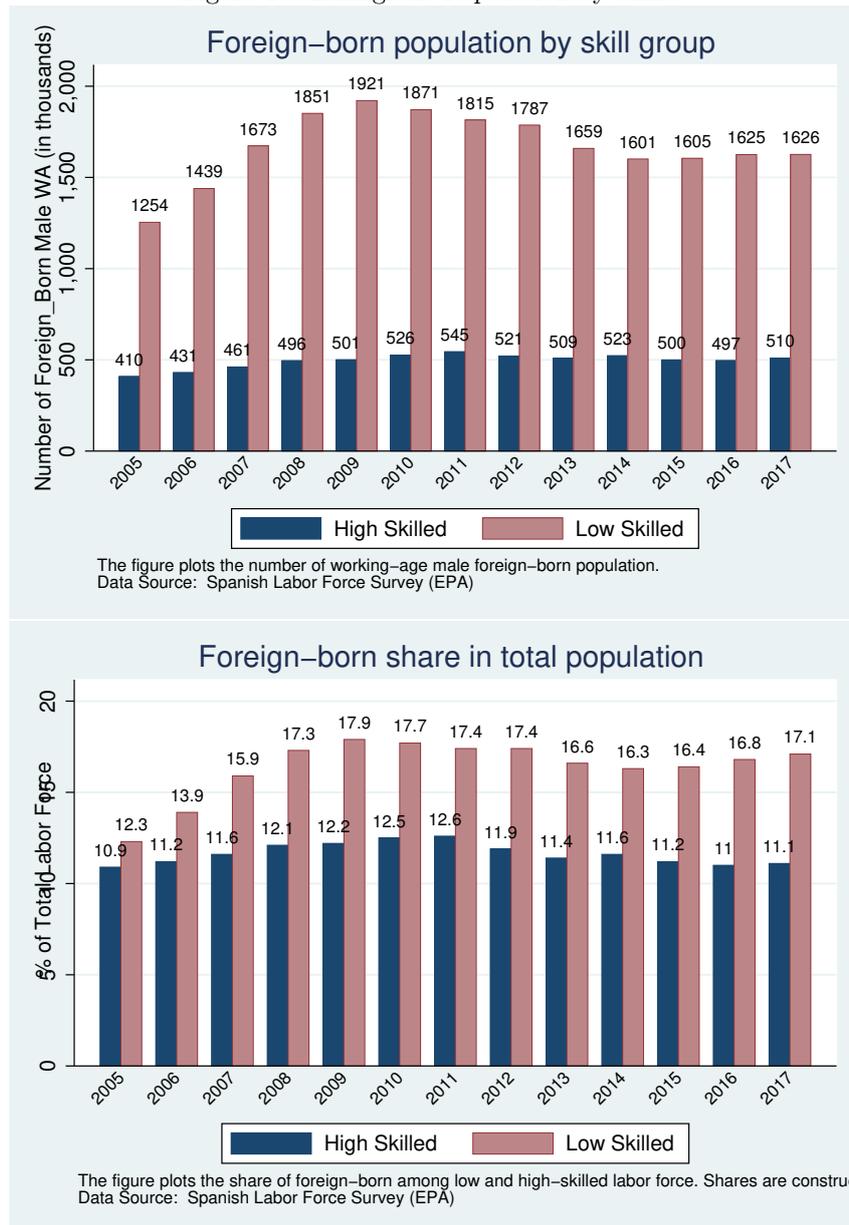
Figure 18: Domestic vs. international departures by Nationality



Immigrant Population by Skill Groups

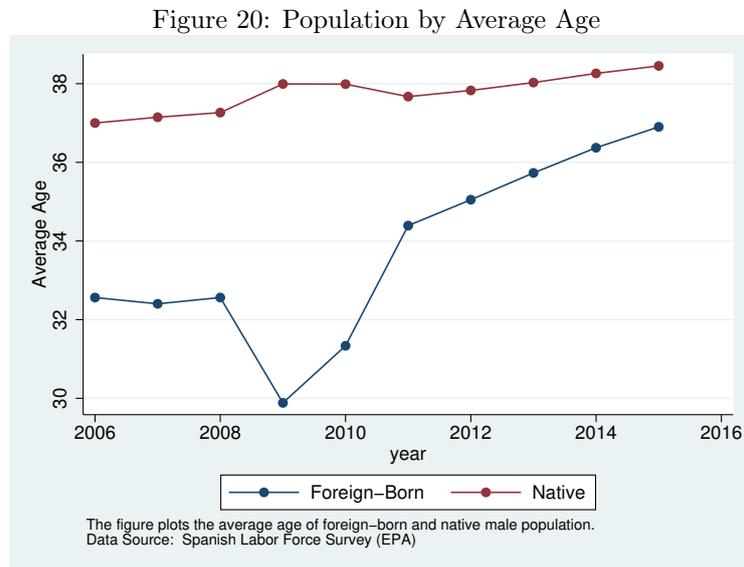
Figures show the increase in the immigrant population stocks. The first graph displays the immigrant population stocks by the level of education while the second shows the same stocks as a share of the total population within that skill group. Individuals with a university degree or higher are considered high skilled, while those with less than a university degree are considered low skilled.

Figure 19: Immigrant Population by Skills



Immigrant population by age groups

Figure show the average age for immigrant and native male population



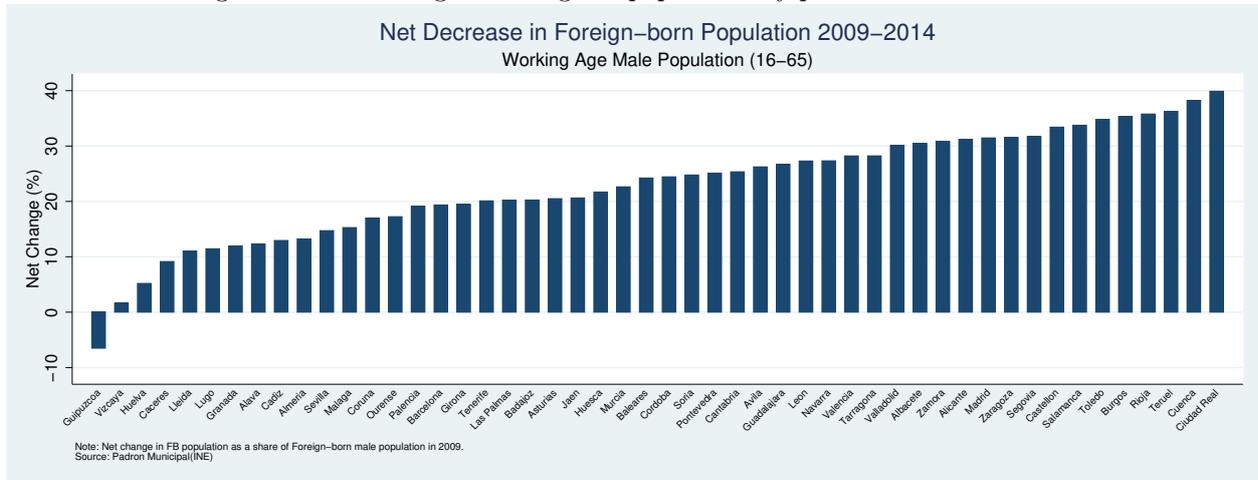
3.13.8 Government Measures

Faced with increasing unemployment, the Spanish government took various measures to reduce the entries and encourage returns of the foreigners. First, the shortage lists that are used for both the Regime Generale, which exempts nominal requests from a labor market test and the *Contingente*, used as criteria for anonymous recruitment from abroad, were curtailed significantly. In October 2008, the quarterly *Regime Generale* “catalogue of difficult-to-cover occupations” contained 32% fewer occupations than the previous list. The occupations which were eliminated, however, represented almost all hiring from abroad. Some occupations (painters, care assistants, waiters, bricklayers, welders, electricians, carpenters, locksmiths, cooks, gardeners, agricultural laborers) disappeared altogether. Only very specific occupations (sports, trainers, doctors, neurosurgeons, dentists, optician, nurses or physiotherapists; specialized mechanics) – mostly qualified – remained. These cutbacks continued in the first list for 2009. The reduction was the sharpest that was observed in the whole OECD countries (OECD, 2009). Second, the ceiling for non-seasonal workers to be recruited anonymously from abroad (the *Contingente*) was reduced (OECD, 2009). In mid-December 2008, the *Contingente*, which sets annual regional caps by occupation for workers was set at 901 for 2009, compared to 15 731 in 2008.

In addition to reducing the inflows, the government also tried to increase outflows in order to support voluntary returns of unemployed, Spain developed the Program for Early Payment of Unemployment Benefits to Foreigners (the *APRE*) in 2008 which allowed third-country nationals to receive on advance an accumulated payment of their unemployment benefits in two lump sums on the condition they return home and do not come back to Spain for at least three years. The proposal provided 40% of the benefit in Spain and 60% upon return and became active in November. Only the 19 countries with bilateral social security agreements are eligible, and the offer is not valid for EU citizens. While the government initially expected many unemployed to apply, uptake has not reached targets. 11 419 unemployed immigrants, mostly from Latin America, signed up for the program by April 2010, while the government calculated that more than 80 000 were eligible. It is, however, difficult to evaluate, at this stage, the full impact of this program even if experience has shown that financial incentives are usually insufficient to drive large return migration.

3.13.9 Net Immigrant Departures as a Share of Total Working-age Immigrant Population

Figure 21: Net Change in immigrant population by province 2009-2016



3.13.10 More on Mobility Data

Padrón is a very useful tool for accounting immigration and internal moves. Regarding the departures however the data suffers from both undercounting and differences in the exact timing of the departure and the deletion from the registers. In this section, I discuss both issues and show that empirically they do not pose a threat to my findings.

Delays in the departures

As discussed earlier, due to lack of incentives, those who leave do not necessarily take the time to de-register. This problem has been partially addressed for the foreigners as, since 2006, individuals who do not renew their registry are deleted from the Padrón registers. Although the deletion process corrects for the departures, it can have a delay of up to two years. Analysis using the long differences in Section 3.8, partially addresses this problem as, rather than focusing on the precise impact of annual departures it covers a longer time period which is longer than the two year window. For instance, in column 3 of Table 12 where regressions are done in three year differences, I still get significant results.

In order to alleviate concerns on the issue further, I provide additional results where I use the average of the departures in $t-1$, t and $t+1$. columns 1 and 3 of Table 25 show the main results for comparison, while columns 2 and 4 present results where the treatment is the average of the three years. It can be seen that elasticities are slightly larger when using the three year averages, and highly significant. More importantly, the measurement issue, and the potential lags in the deletions do not bias my results.

Table 25: Robustness: Three Year Average Treatment, 2009-2014

	Wage		Employment	
	2SLS	2SLS	2SLS	2SLS
Net Outflow Rate	2.0286 (0.609)***		2.4122 (0.752)***	
Three year average		2.6842 (0.737)***		3.1484 (1.078)***
N	250	250	250	250
KP F-Stat	17.45	8.27	17.45	8.27

The table reports 2SLS estimates for the impact of net change in immigrant population in an area. The Net Outflow Rates in columns 1 and 3 are measured as the decrease in the number of working-age male immigrant population between t and $t - 1$ as of total working-age male population in year $t - 1$. Net Outflow Rates in columns 2 and 3 are measured as the average of the decrease in across three years. Rate in period t is the average departure rate of period $t - 1$, t and $t + 1$. Regressions are estimated annually, across 50 provinces over 5 periods (2009-2014). Regressions are weighted by total employment in the base year. Standard errors are clustered at province-level.

Data source: MCVL and INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Measuring Emigration of Spanish nationals

Measurement of the emigration of the Spanish citizens can be particularly problematic as they are not required to renew their inscription in the Padrón. A national will be removed from Padrón, only if he or she registers at a Spanish consulate abroad, in which case the person would be included in the Register of Spaniards Living Abroad, or PERE (*Padrón de Espanoles Residentes en el Extranjero*, in Spanish). Many Spaniards living in another country do not register in a consulate because it conveys little or no advantage. On the other hand, those who are removed from the Padrón lose their health care and other subsidies or benefits, and many find it harder to exercise their voting rights in Spain (Arango (2016); González-Ferrer and Moreno-Fuentes (2017)). Furthermore registering in a consulate may not be convenient, as it requires producing an official document attesting that the emigrant will stay abroad for more than a year. Due to these issues, it is likely that both Padrón and the PERE underestimate the number of Spanish emigrants, particularly the native born.

This measurement issue of the native outflows may not be of an empirical issue for a few reasons: First, I am using the native-born outflows as a control variable. Since I am not interested in the precise magnitude of the coefficient, the overestimation of that particular coefficient is not problematic. Second, it includes departures abroad but also to other provinces. Although Padrón mismeasures the emigration, it is very precise for internal flows. Given that natives' low rates of emigration, this reduces the importance of the issue even further. Third, I am using a differences-in-differences estimation strategy comparing the relative changes across provinces. Given that the mismeasurement issue is a valid concern for all of the provinces, it should not bias the estimated coefficients. Furthermore, I am taking first differences which absorbs any province-specific fixed characteristics. If for some reasons, certain provinces systematically under/over-report native departures, it should be taken care in the first differences.¹⁰⁵

¹⁰⁵The central government determines the level of local fundings in municipalities based on the population stocks using Padrón. This can incite local authorities to inflate the population numbers in order to increase their funding.

3.13.11 Construction of the instrument

Predicting Native Stocks

As discussed in the Section 3.5.2, stock of natives residing in a given province might depend on the number of immigrant in the same location or on unobservables correlated with the labor market outcomes. To address the issue, I use a shift share strategy to predict the location choice of natives, based on their past distribution across Spain.

Following a strategy similar to Sanchis-Guarner (2017), I use the past distribution of natives based on their province of birth (the “share”), and distribute the current population (the “shift”) accordingly.

The strenght of the instrument depends on the historical mobility of natives from different provinces and the strength of ethnic networks. Some regions have historically had larger mobility propensities (Galicia), and some bilateral internal migration flows are based on historical location patterns (for example Galicians in Madrid or Andalusians in Cataluña). A person born in a given province b can either stay where he/she was born (stayers) or can move and reside in a different province $j \neq b$ (movers). R is the total number of provinces in Spain in which natives can locate. I need individual level data to know the province of birth. For this purpose I use native location patterns from census 1991 as base year. I define the share of stayers in province j as the proportion of natives born and living in a province over all the natives born in the province (regardless of where they reside) in 1991. In this case, the province of birth and residence is the same, i.e $j = b$. The stayers share is defined as follows:

$$share_{j(j=b),1991}^b = \frac{natives_{j=b,1991}^b}{\sum_j^R natives_{j,1991}^b} \quad (31)$$

To obtain yearly predictions of the number of immigrants by nationality n , I multiply expression 31 by annual national stock of natives $natives_t^b$ of province b living in Spain in period t . The imputed native stock of a specific province b in province j at time t is thus calculated allocating yearly total stocks weighted by its historical share (3.7.3):

$$\widehat{natives}_{j,t}^b = (natives_{Spain,t}^b) * share_{j,1991}^b \quad (32)$$

To calculate the imputed total native stock in province j at time t , I sum (32) across provinces of origins (R):

$$\widehat{natives}_{j,t} = \sum_n^R (\widehat{natives}_{j,t}^b) \quad (33)$$

Predicting National Stocks of Immigrants: Gravity Equation

The instrument's validity is conditional on immigrant outflows and labor market outcomes of natives being orthogonal to the local economic shocks. Given that the national shock (which is the sum of local shocks) partially determines the national stock of immigrants in Spain, this can create concerns about endogeneity. To address this concern, I construct instruments where I use national stocks that are predicted through a gravity equation that is based on factors that are specific to the country of origin (push factors) that are plausibly exogenous to local shocks in Spain.¹⁰⁶

I follow a similar strategy to Saiz (2007); Ortega and Peri (2009); Sanchis-Guarner (2017); Bahar and Rapoport (2018) and compute the yearly predicted total stock of immigrants by country of origin from the results of a gravity model which depends only on push factors which are specific to the country of origin of the immigrants. Specifically I follow Frankel and Romer (2019), I use a gravity-type model that only contains push-factors from origin to predict the total stocks from nationality n in Spain in a given year t .¹⁰⁷ The estimated equation is:

$$\ln(\text{Foreign}_{Spain,t}^n) = \chi' \ln(\text{Economic}_{n,t-1}) + \psi' \ln(\text{Geographic}_{n,t-1}) + \eta_g + \vartheta_t + v_{n,t} \quad (34)$$

where $\text{Economic}_{n,t-1}$ is a matrix of (lagged) time-varying economic conditions of the sending country (log of gross domestic output in real terms, square of GDP, log of total population, percentage of urban population, log of used agricultural area, value added of services, value added of industry, average years of life expectancy, unemployment rate and a dummy of belonging to the EU27). $\text{Geographic}_{n,t-1}$ is a matrix of time-invariant geographic characteristics of the sending country (log of distance to Spain, log of area, number of cities, latitude and longitude and dummies for common language, common border and common colonial past with Spain). I include year dummies ϑ_t and country-group dummies η_g . I can alternatively include country dummies, which drops the time-invariant variables. The economic and country specific variables come from the World Bank, the geographical and distance variables come from Centre d'Études Prospectives et d'Informations Internationales (CEPII).

I run this regression using data for 109 countries, which represent more than 98% of the immigrant stock into Spain for the period. Results for different specifications are showed in table 26. Column titles correspond to the log transformations. In columns 1 and 2 include country fixed-effects while column 3 includes only region fixed effects. All the models have high predictive power.

From the results 1 in Table 26 I recover the predicted stocks of immigrant in Spain from nationality n for every year 1988-2017. I use the prediction from estimates from column 1 for the construction of the

¹⁰⁶The endogeneity of national shocks to the immigrant stocks at local units is especially a valid concern if the spatial units are small and the economic conditions that attract immigrants are spatially correlated. For instance, the economic condition in large provinces such as Madrid or Barcelona could influence the total number of immigrants deciding to come to Spain, even if they end up locating somewhere else in the country based on their ethnic networks.

¹⁰⁷Some examples of other studies that use a gravity model to instrument for migration stocks are Saiz (2007); Felbermayr et al. (2010); Ortega and Peri (2014); Alesina et al. (2016); Sanchis-Guarner (2017); Bahar and Rapoport (2018).

instrument, and I use the rest of the specifications estimates for the robustness checks. These predicted stocks replace annual national stock of immigrants ($Foreign_{j,t}^n$) term in equation 15.

Table 26: Gravity Equation: Predicting National Stocks

	Log(Male)	asinh(Male)	Log(Male)
L.Log GDP	1.0429 (0.330)***	1.0428 (0.330)***	1.1833 (0.607)*
L.Log Population	-1.3320 (5.140)	-1.3337 (5.139)	-33.0600 (12.295)***
L.Log Urban Pop	-4.1464 (9.278)	-4.1448 (9.278)	-15.7568 (18.014)
Log of GDP squared	0.5235 (0.175)***	0.5235 (0.175)***	0.1973 (0.422)
Dummy for EU27=1	546.3180 (4949025.461)	546.3125 (4948817.125)	-0.0668 (0.381)
L.Log Land Area	2.8137 (8.567)	2.8132 (8.567)	1.2796 (13.352)
L.Life Expectancy	0.0037 (0.033)	0.0037 (0.033)	0.0119 (0.033)
L.Agricultural Land	-0.0128 (0.013)	-0.0128 (0.013)	-0.0021 (0.005)
L.Services, VA	0.0000 (0.000)	0.0000 (0.000)	0.0000 (0.000)
L.Industry, VA	-0.0130 (0.010)	-0.0130 (0.010)	0.0083 (0.011)
L.Unemployment Rate	0.0107 (0.008)	0.0107 (0.008)	0.0219 (0.022)
L2.Log GDP	-1.4668 (0.450)***	-1.4668 (0.450)***	-2.3342 (0.778)***
L2.Log Population	-0.8043 (5.159)	-0.8026 (5.158)	34.3745 (12.272)***
L2.Log Urban Pop	5.3849 (9.266)	5.3834 (9.266)	17.9993 (17.857)
L2.Log Land Area	15.4207 (5.115)***	15.4203 (5.115)***	-1.1598 (13.339)
Common Official Primary Lang=1	0.0000 (.)	0.0000 (.)	0.9277 (0.654)
Language Spoken=1	0.0000 (.)	0.0000 (.)	1.5103 (0.305)***
Common Colonizer=0	0.0000 (.)	0.0000 (.)	0.0000 (.)
Log Distance Capital	0.0000 (.)	0.0000 (.)	-2.5575 (0.310)***
N	1889	1889	1889
KP F-Stat			

The unit of observations are country of origins (109), period is 2000-2017. Errors are clustered at the country level. columns 1 and 2 include country fixed effects, while column 3 includes only region fixed effects. Data source: CEPII, INE, World Bank.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.13.12 Instrument Validity

Exogeneity of the base year

In constructing the instrument, I use the distribution of the immigrant population in 1991 as the “share”. Instrumental variables estimation will be consistent if the geographical location patterns of the 1991 stock of immigrants are uncorrelated with province-specific shocks that affect the labor market outcomes of the natives between 2009 and 2016. This base year is 18 years before the main period of analysis, which provides a substantial amount of lag between the two. Still, it is possible that some unobservable shocks that determined the distribution of the immigrant population in the base year were still present and thus continued to impact the evolution of the labor market outcomes of the natives during the period of analysis despite the inclusion of time and province fixed effects.

In order to address this concern, I follow Gonzalez et al. (2009) and Sanchis-Guarner (2017), and I regress the provincial share of immigrant population in 1991 on that year’s economic conditions and, then the change in this share during my observation period (2002-2016) on these same variables. The aim of this exercise is to show that the determinants of the geographical distribution of the mass of immigrant workers in 1991 does not perfectly predict the location during my period of analysis. The results are shown in Table 27.

Table 27: IV validity checks: Base-year validity regressions

	FB Share 1991	Change 2002-2016
Log Disposable Income	-0.0620 (0.027)**	-0.0540 (0.038)
Log GDP	0.0608 (0.027)**	0.0460 (0.038)
Average Daily Wage	-0.0008 (0.000)**	0.0012 (0.001)**
Share of Agriculture in VA	-0.1003 (0.041)**	0.0450 (0.059)
Share of Industry in VA	0.0216 (0.029)	-0.0141 (0.041)
Share of Services in VA	0.0940 (0.028)***	0.0147 (0.039)
Foreign-born Unemployment Rate	-0.0614 (0.026)**	-0.0315 (0.038)
Constant	0.0076 (0.041)	0.0661 (0.058)
N	50	50
Adj R2	0.57	0.33

The unit of observations are provinces. All the regressors are from year 1991. The omitted category is share of construction in value added (VA). Data source: INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In the first column, I present results where the dependent variable is the share of the immigrant population in 1991. I include the log of disposable income, the log of GDP of the province, the log of average daily wage, the share of different sectors (agriculture, services, and industry) in the regional value-added (excluding the construction sector), the unemployment rate for immigrant workers. The regression includes the 50 provinces observations and a constant, so all the values are relative to the national mean. The model has high predictive power, and most of the regressors are significant, showing that economic factors influenced the location decision of the immigrant in 1991.

In the second column, I regress this same set of variables on the change of immigrant share over 2002-2016, which is the whole period of analysis. Apart from the average daily wage, all other variables are statistically insignificant. This test supports that using 1991 as the base year is appropriate. Moreover, given that the analysis is done in first-differences, most of these 1991 conditions are netted-out.

Overlapping response problem

Instrumental variables based on the shift-share methodology have been subject to a fair amount of discussion and criticism (e.g. Jaeger et al., 2018; Goldsmith-Pinkham et al., 2018). As discussed by

Jaeger et al. (2018), if local conditions that influence immigrants location decisions are persistent, then the exclusion restriction assumption of the instrumental approach may not be satisfied as the the local labor market adjustments to an immigration-induced supply shock can take time. As discussed earlier, “overlapping response problem” may be less concerning in Spain, given the changes in the country-origin mix that has been observed since 1990s.

Still, in order to address concerned raised in Jaeger et al. (2018), in Table 28, I test the ability of my instrument to deal with the overlapping response problem by including both the contemporary and lagged predicted out-migration flow in the regression. I construct both the contemporary and lagged instruments using the same base period year (1991), so that the variation between the two are driven by national-level changes in the composition of immigration across periods. If the composition of immigrant population does not change sufficiently, both instruments should be highly correlated and both might predict the actual inflows.

Table 28: Multiple Instrumentation, 2009-2016

	Wage			Employment		
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Net Outflow Rate	2.0286 (0.609)***	2.0286 (0.694)***	2.3305 (1.085)**	2.4122 (0.752)***	2.4122 (0.765)***	0.4776 (0.877)
5-year Lagged Outflows		0.7332 (0.257)***			-0.2581 (0.460)	
10-year Lagged Outflows			0.4472 (0.338)			-1.1715 (0.306)***
N	250	250	200	250	250	200
r2	0.26	0.27	0.21	0.49	0.48	0.58
Cragg-Donalds Stat	34.73	17.50	11.77	34.73	17.50	11.77

The table reports 2SLS estimates for the impact of net change in immigrant population in an area, measured as the decrease in the number of working-age male immigrant population between t and $t - 1$ as of total working-age male population in year $t - 1$, on native local wage and employment growth in the aggregate (skilled and unskilled). Regressions are estimated annually, across 50 provinces over 5 periods (2009-2014). Regressions are weighted by total employment in the base year. Standard errors are clustered at province-level.

Data source: MCVL and INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In Table 28, I present results, including only the contemporary out-migration rate (columns 1 and 3) for comparison. In columns 2 and 4, I also include a 5-year lagged instrument as such a lag may capture the dynamics adjustments to regional labor supply shocks. I find that the inclusion of lagged instruments does not change the point estimate of the contemporary out-migration flow both for wage and employment growth. Results in column 2 show that although lagged outflows improve wage growth, it does not change the coefficient size of the contemporary outflows, which are also much stronger predictors. In column 4, I show that lagged outflows do not seem to be statistically significant for employment.

3.13.13 Construction of Bartik control

The Great Depression had differential severity across Spanish provinces. The decline in economic activity in several specific industries, the sectoral composition of provinces might explain an important part of the growth in employment. Following the literature, I build a Bartik control to proxy for industry-driven local demand shocks.¹⁰⁸ The Bartik will absorb local variation in employment resulting from national level changes of sectors which are strongly represented in a particular province. For instance, when employment in a given industry increases (decreases) nationally, areas in which that industry represented a significant share of employment must have experienced a positive (negative) relative change in the demand for workers relative to those where that industry is not present. The predicted growth of local employment, assuming employment in each industry i grows in line with the national rate.

Specifically, I multiply the province level sectoral employment shares in 2005 by the employment level of the sector at the country level in each year t .¹⁰⁹ Following Autor and Duggan(2003) and Goldsmith-Pinkham et al. (2018), I leave-out the employment of the own area, to address concerns that the introduction of own-area employment may mechanically increase the predictive power of the shock. Specifically, I calculate:

$$Bartik_{j,t} = \sum_i \phi_{j,t_0}^i L_{i(-j)t} \quad (35)$$

where ϕ_{j,t_0}^i is the share of employed individuals in area j , at time the start period (t_0) working in a two-digit industry i (53 sectors) . $L_{i,(-j),t}$ is the number of workers employed in industry i at time t nationally, excluding area j . The Bartik instrument predicts the level of employment in a province, if the local industry shares had remained the same as in the starting year and employment had grown in local firms at the same rate as in same-industry firms in the rest of the country.¹¹⁰ Given that my specification is at first differences, I take the difference in the predicted employment levels to obtain my control.

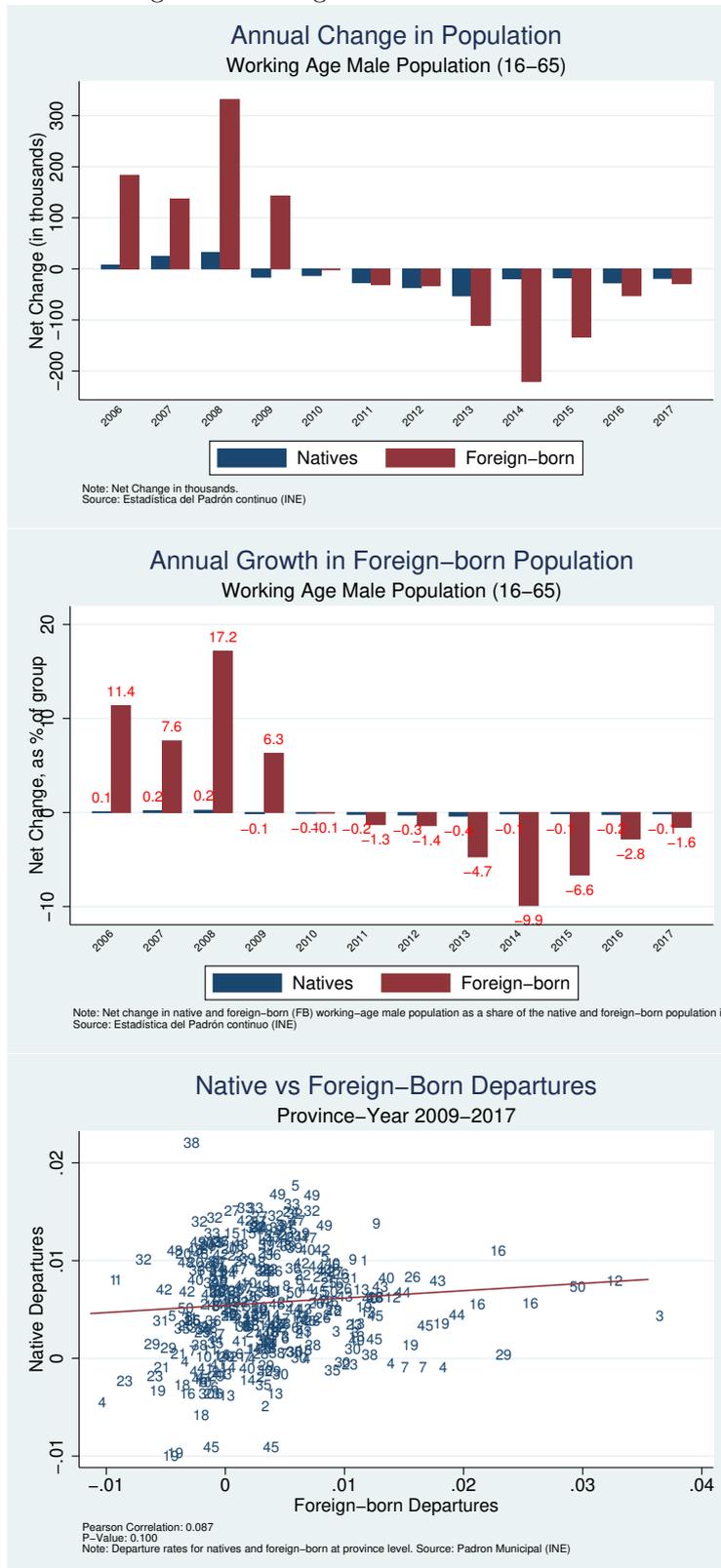
¹⁰⁸See for instance Cadena and Kovak (2016); Basso et al. (2018) for a use as an instrument, and Lee et al. (2017) as a control or the main regressor.

¹⁰⁹I build my Bartik control to obtain the predicted employment levels, similar to Basso et al. (2018). Another option is to calculate the change in sectoral employment at national level and then distribute the change across provinces based on the initial distribution (see Cadena and Kovak (2016); Lee et al. (2017) for instance). I prefer the first option as, given the first differences approach, the interpretation is easier. However, I test the robustness of my results using alternative measures. I get similar results.

¹¹⁰Using earlier (e.g. 2000) or later (e.g. 2009) base years does not change my results.

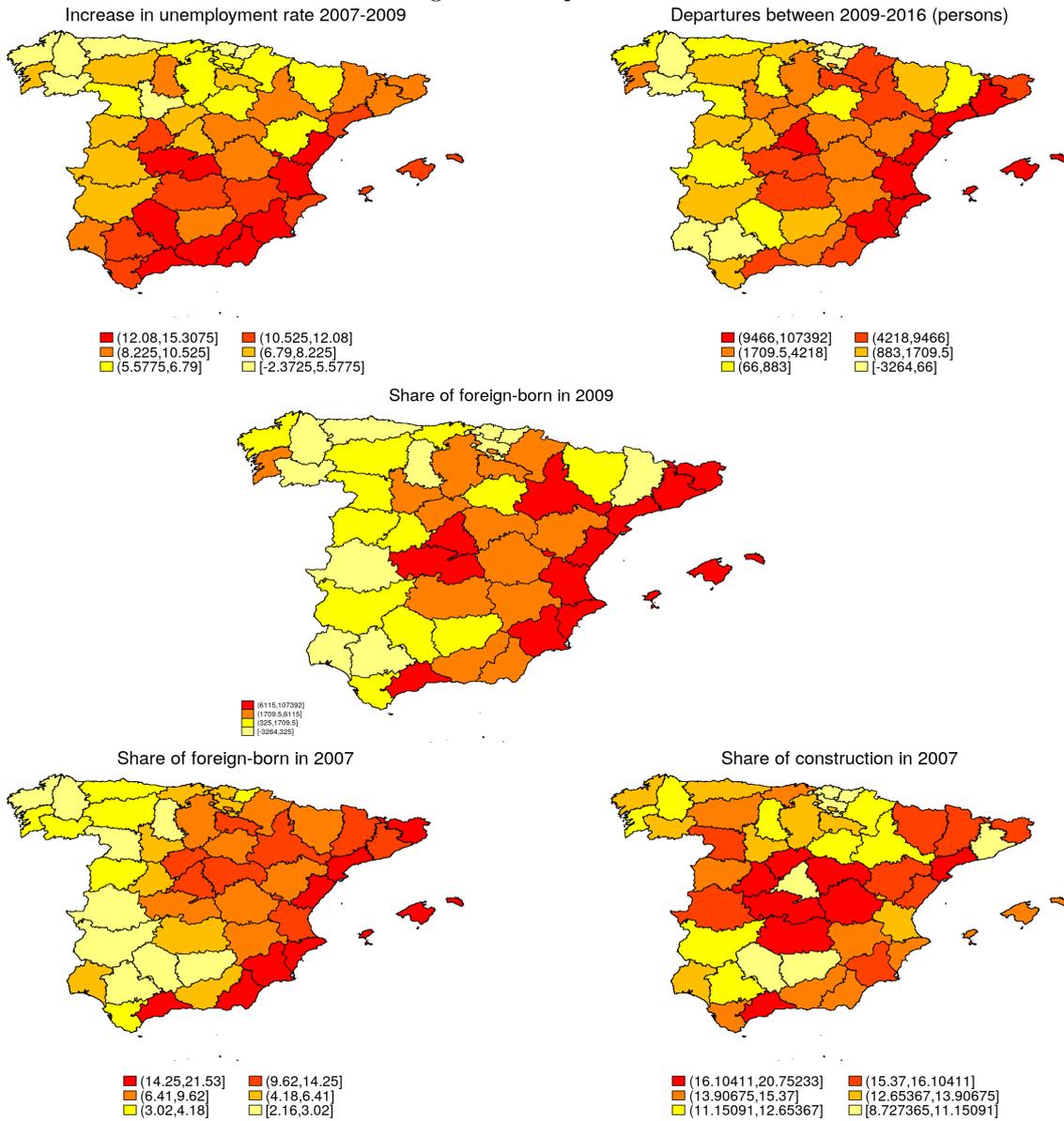
3.13.14 Immigrant and Native Outflows

Figure 22: Immigrant and native outflows



3.13.15 Maps

Figure 23: Maps



3.13.16 Differentiating Between Boom and Bust

Results in the paper show that the out-migration of the immigrant accelerated both the wage and employment growth of the natives during the Recession. In this section, I extend the analysis to the previous years - the boom years - to see whether symmetric effects can be found during an economic expansion. In Table 29 I use the exact same specification for for the boom years (2002-2007) and for the bust - or Recession - years (2009-2014). For comparability with the results from the Recession, I define the boom period as the 5 years prior to the crisis to match the length of the Recession period analysis.¹¹¹

Table 29: Boom vs. Bust

Panel A: Change Wages				
	2002-2007	2002-2007	2009-2014	2009-2014
Net Outflow Rate	0.2180 (0.156)	0.2217 (0.154)	2.0286 (0.609) ^{***}	1.9905 (0.699) ^{***}
Bartik		0.0200 (0.056)		0.0474 (0.190)
N	250	250	250	250
KP F-Stat	20.50	20.20	17.45	16.64
Panel B: Change Employment				
	2002-2007	2002-2007	2009-2014	2009-2014
Net Outflow Rate	1.7525 (0.530) ^{***}	1.8529 (0.507) ^{***}	2.4122 (0.752) ^{***}	1.6846 (0.790) ^{**}
Bartik		0.5413 (0.117) ^{***}		0.9048 (0.145) ^{***}
N	250	250	250	250
KP F-Stat	20.50	20.20	17.45	16.64

The table reports 2SLS estimates for the impact of net change in immigrant population in an area, measured as the decrease in the number of working-age male immigrant population between t and $t-1$ as of total working-age male population in year $t-1$, on native local wage and employment growth in the aggregate (skilled and unskilled). Regressions are estimated at the yearly level, across 50 provinces. All regressions are weighted by total employment in the base year include year fixed effects. Standard errors are clustered at province-level.

Data source: MCVL and INE.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel A and Panel B present results for wages and employment, respectively. Columns 1 and 2 show results for the pre-crisis period, columns 3 and 4 show those for the Recession period. In columns 1 and 3, I regress the labor market outcomes on the net outflow rate. Panel A shows that the magnitude of the coefficient is almost ten-folds larger during the Recession than in the growth period, although it is statistically insignificant. In columns 2 and 4, I add Bartik to control for differences in the growth rates between two periods. The results do not change.

In Panel B, I repeat the exercise for the employment margin. In columns 1 and 3, I regress the employment growth on the net outflow rate. During the growth period, an increase of 1% in the out-

¹¹¹I check whether the outcomes for the pre-crisis period are sensitive to the definition of the period. The results available upon request.

migration rate accelerated employment growth of the native-born by 1%, which is two times smaller than the 2% which is observed during the Recession. This suggests that net outflows have a stronger effect during economic busts than economic growth periods. In columns 2 and 4, I add Bartik to control for differences in the growth rates between the two periods. Reassuringly, once I control for the differences in the labor demand, the coefficients for both periods are similar.

The findings for the Recession is in line with the literature that argues that immigrant mobility can have a cushioning effect for the natives during a demand shock Amior (2017); Cadena and Kovak (2016); Basso et al. (2018). The results for the growth period on the otherhand are novel.

3.13.17 Part-time Employment

The results presented in the paper focus on full-time employment. The outflows can also increase part-time employment and impact the wages of those who hold such jobs. In order to test this, I repeat the main specification but using part-time employment as the outcome variable. Table 30 presents these results. columns 1-2 show that the outflows did not increase the average wages of part-time jobs. However, larger outflows increased the number of natives holding part-time jobs, for both skill groups.

Table 30: Part-time Employment and Wages

	Wages		Employment	
	Low Skilled	High Skilled	Low Skilled	High Skilled
Net Outflow Rate	0.2269 (0.321)	0.3388 (0.842)	2.8739 (1.605)*	5.5630 (3.034)*
N	250	247	250	250
KP F-Stat	21.87	18.25	21.87	18.27

The wages and employment growth in the previous periods, standardized by initial employment. The table reports OLS and 2SLS estimates for the impact of net change in immigrant population in an area, measured as the decrease in the number of working-age male immigrant population between t and $t - 1$ as of total working-age male population in year $t - 1$, on native local part-time wage and employment growth in the aggregate. The unit of observations are provinces. Regressions are weighted by group-specific employment in year $t - 1$ and year fixed-effects. Standard errors are clustered at province-level. Data source: MCVL and INE

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4 Migration and Post-conflict Reconstruction: The Effect of Returning Refugees on Export Performance in the Former Yugoslavia

Abstract

During the early 1990s Germany offered temporary protection to over 700,000 Yugoslavian refugees fleeing war. By 2000, many had been repatriated. We exploit this natural experiment to investigate the role of migrants in post-conflict reconstruction in the former Yugoslavia, using exports as outcome. Using confidential social security data to capture intensity of refugee workers to German industries—and exogenous allocation rules for asylum seekers within Germany as instrument—we find an elasticity of exports to return migration between 0.08 to 0.24. Our results are stronger in knowledge-intensive industries and for workers in occupations intensive in analytical and managerial skills.

Keywords: migration, refugees, knowledge diffusion, management, exports, productivity

JEL Classification Numbers: O33, F14, F22

4.1 Introduction

What do the emergence of the textile sector in Prussia, of the IT sector in India or Israel, of the garment industry in Bangladesh, or of the car-parts industry in Bosnia have in common? One particular aspect stands out: the circumstances of their birth and success can be traced back to a migration episode of some sort; and in all cases, the available evidence –be it anecdotal or empirical– points to migration-driven transfers of ideas, knowledge and technology as the driving force behind them.¹¹² For the most part, the economic debate on immigration has focused on its short-term labor market and fiscal effects. Perhaps because of this, less attention has been given to the long-run economic opportunities linked to migration. Nevertheless, there is a robust and growing literature documenting many aspects through which migration contributes to long-run growth, such as through innovation (e.g. Kerr, 2008; Choudhury, 2016), through skill-complementarities arising from diversity (e.g., Ortega and Peri, 2014; Alesina et al., 2016), or through the reduction of bilateral transaction costs resulting in higher trade (e.g., Gould, 1994; Rauch, 1999; Rauch and Trindade, 2002; Parsons and Vézina, 2018b) and investment flows (e.g., Kugler and Rapoport, 2007; Javorcik et al., 2011; Kugler et al., 2017) across borders. In this paper we explore a novel and additional angle: the role that migrants play in spreading ideas, technology and knowledge across countries, and how such transfers are reflected in long-term real economic outcomes such as exports.¹¹³ In particular, in this paper we exploit a natural experiment and document how return migrants –having spent time in a foreign country– explain the subsequent performance of the same export sectors in which they had worked while abroad.

The context of our study is the early 1990s, when about 700 thousand citizens of the former Yugoslavia fled to Germany escaping war. Most of the Yugoslavian migrants in the first half of the 1990s were given a *Duldung* status (German for “toleration”), in effect a temporary protection status, or more specifically, a “suspended deportation” permit. After the Dayton peace agreements were signed in 1995, the protection status of temporary migrants was revoked, forcing them to leave the country. By 2000, the majority of these migrants had been repatriated back to their home country or to other territories of the dissolved Yugoslavia. We exploit the stay of these refugees in Germany, and the subsequent massive inflow of return migrants –with experience in the German workforce– into the former Yugoslavia, to study sector-specific productivity shifts as reflected by export performance.¹¹⁴ To do so we rely on confidential administrative data from the German Institute for Employment Research (IAB), which we use to compute the number of Yugoslavian migrants working in a particular 4-digit industry who had arrived to Germany during the

¹¹²For the textile sector in Prussia see Hornung (2014); for the IT sector in India and in Israel see Khanna and Morales (2017) and Rosenberg (2018), respectively; for the garment industry in Bangladesh see Rhee and Belot (1990) and Easterly (2001).

¹¹³Other studies looking at the role of foreign workers inducing productivity shifts within-plants and within-firms include Markusen and Trofimenko (2009) and Poole (2013a), respectively, using native workers’ wages as main outcome. In previous work, we study the link between migration and productivity in cross-country comparisons, using the emergence and growth of export sectors as main outcomes (Bahar and Rapoport, 2018).

¹¹⁴Following Bahar et al. (2014) and Bahar and Rapoport (2018) we use changes in exports for a particular product as a proxy for productivity improvements.

Balkan refugee crisis and returned home after the war. We link this information to standard disaggregated international trade data and employ a difference-in-differences methodology to estimate changes in export values from Yugoslavian countries to the rest of the world caused by return migration of workers who were employed in those same sectors in Germany. In order to address concerns of endogeneity due to, for example, self-selection of workers into certain industries with high potential back home, we instrument the actual number of returning workers per industry with their expected number given a spatial dispersal policy that exogenously allocated asylum seekers across the different regions of Germany upon arrival.

We find that, on average, industries with a one percent increase in return migration experienced an increase in exports to the rest of the world of 0.08 to 0.24 percent between the pre and post-war periods. In fact, the estimated elasticity keeps increasing as time passes after refugees have returned. We also show that our results cannot be explained by an existing previous trends on exports or are driven by mere convergence between the industry structure of the former Yugoslavia in the 2000s and that of Germany in the 1990s. In a robustness section, we perform a number of tests to rule out plausible alternative explanations, such as investment linked to migration or a decrease in information costs linked to international trade due to migrant networks; we also focus on the specific case of Bosnia, for which we are able to investigate economic outcomes beyond exports, and find that returning refugees positively affect the number of firms and employment levels in an industry, too.

The last part of the paper explores candidate mechanisms behind our results. We find evidence consistent with the idea that migrant workers exposed to industries in Germany bring back knowhow, knowledge and technologies that translates into higher productivity in those same industries, which in turn is reflected in export performance. In particular, we find that our results are driven by knowledge-intensive industries, as well as by certain types of workers and occupations which would be more suited for diffusing productivity-inducing knowhow across borders. For instance, our results are driven by workers with high educational attainment, in occupations intensive in analytical tasks (as opposed to manual ones), occupations that can be classified as professional and/or skill-intensive, and occupations that have managerial characteristics. We also find that our results are stronger when looking at workers who, while abroad, experienced fast wage growth, and were employed by the top paying firms within each industry. As such, our paper belongs to the burgeoning literature that looks at foreigners and migrant returnees as drivers of knowledge diffusion (e.g. Kerr, 2008; Markusen and Trefl, 2009; Poole, 2013a; Choudhury, 2016; Hausmann and Neffke, 2016; Malchow-Møller et al., 2017; Bahar and Rapoport, 2018; Kerr and Kerr, 2018; Giorcelli, 2019), as the transmission of tacit or non-codifiable knowledge requires human interaction (Arrow, 1969; Polanyi, 1966), as well as to the growing literature emphasizing the role of management as a crucial determinant of productivity (e.g., Bloom and Van Reenen, 2007; Bloom et al., 2012, 2013, 2019). Our results are consistent, too, with recent work by

Bloom et al. (2018) who find that management has a positive effect on the capacity of firms to export.¹¹⁵

One valid concern is whether our results could be explained by convergence patterns reflected in structural transformation processes (e.g., the export basket of post-war Yugoslavia converging towards the export structure of pre-war Germany), or to scale effects (more workers returning mechanically boost exports due to an increase of the supply of industry-specific labor). We have enough evidence to rule out those possibility. First, with respect to convergence, our main specification includes a term that controls for convergence effects, given the pre-war structure of exports in Germany. We also perform a number of falsification tests to show that our results are particular to Yugoslavia, and not part of a trend in countries with similar export structure at the baseline period. Finally, we show that our results are driven by within-industry variation based on characteristics of the worker returnees and of their particular occupations, as we show in the last section of the paper. In that sense, there is no reason that convergence would occur only for certain industries and not for others, in ways that also explain those within-industry distributions. Secod, when it comes to scale effects, we also present evidence that rules out this possibility. We do so by showing that our effect is driven by both labor- and capital-intensive goods, showing no signs that our results are capturing a traditional Rybczynski effect. We also show that our results also hold for a constructed measure of export productivity based on Costinot et al. (2012), besides export value. Finally, as explained before, we show our results are driven by certain type of workers such as those with occupations intensive in analytical skills, or with managerial characteristics, which typically corresponds to just a small number of workers in each industry.

This paper contributes to the literature in international economics and economic development in several ways. First, to the best of our knowledge, it is the first study that uses a natural experiment as a source of identification to causally estimate the effect of migration flows on country-wide and sector-specific real economic outcomes such as export performance and, more broadly, productivity shifts. Our findings suggest that migrants play a role in the development of nations by diffusing knowledge and technologies across countries.¹¹⁶ Second, we contribute to the economic growth and development literature by exploring the role of migrants in the process of structural transformation, which is evidenced to correlate with economic stability and growth (e.g., Imbs and Wacziarg, 2003; Hausmann et al., 2006; Hidalgo et al., 2007; Koren and Tenreyro, 2007; Krishna and Levchenko, 2013; Cadot et al., 2011). To the best of our knowledge, we are first to do so in the context of forced migration (for more on this

¹¹⁵Our results are also consistent with plenty of anecdotal evidence we collected for the case of Yugoslavian refugees. For example, the story of Suad Bešlić. Bešlić was a soldier in the Bosnian forces during the Balkan wars in the early 1990s who arrived to Germany heavily wounded in 1994. After a long recovery (which allowed him to extend his stay), he followed studies in Germany in car design. After graduation, he started working as a designer in Lenthmer, a leading manufacturer specialized in firefighting and rescue vehicles. Over the years, he became one of the top designers of the company. Upon his return to Bosnia, Beslic set up his company which produces and exports parts used in modern fire trucks. All of his workers go through training both in-house and in Germany. For other examples see Online Appendix Section 4.10.1.

¹¹⁶An extreme case of human-driven technological knowledge diffusion is the case of industrial espionage. For example, Glitz and Meyersson (2017) show that industrial espionage was a channel for knowledge diffusion between West and East Germany (the GDR) during the Cold War. The sectoral information provided by informants working across West German industries, helped East Germany increase its sectoral productivity and narrow West-to-East TFP gaps. They further show that the effects were particularly strong if the information originated from the most advanced sectors.

literature, see Becker and Ferrara, 2019). In that sense, our paper documents a historical episode where refugees played an important role in the post-conflict economic reconstruction of their home country upon returning. Our results, therefore, also speak to the post-conflict reconstruction literature in general, and in the former Yugoslavia more specifically (Black, 2001; Black and Gent, 2006).

The rest of the paper is divided as follows. Section 4.2 provides a description of the historical context of the Yugoslavian refugee crisis. Section 4.5 details the data sources. Section 4.6 explains the setting and the empirical strategy. Section 5.7 presents the main results and performs a series of robustness tests. Section 4.8 explores differential results based on types of migrant workers' characteristics and their occupations. Section 4.9 concludes.¹¹⁷

4.2 Historical Context

In June 1991 the Socialist Federal Republic of Yugoslavia started to disintegrate following several armed conflicts and ethnic civil wars. Fighting began with the "Ten-Day War" in the summer of 1991 after Slovenia declared its independence. Soon thereafter the conflict spread to Croatia and later on, in 1992, to Bosnia and Herzegovina. It was only in December of 1995, upon the signing of the Dayton Peace Accord involving President Clinton, that the armed conflict officially ended.¹¹⁸

During the armed conflict, around 3.7 million people (roughly 16 percent of the Yugoslavian population) were displaced and fled from their homes, making this episode the largest migration flow in Europe since the end of the Second World War (Radović et al., 2005). While many affected by the war became internally displaced, about 800 thousand people resettled outside of the boundaries of the former Yugoslavia, hoping to find refuge in foreign countries (Lederer, 1997).¹¹⁹ Among these countries, Germany was one of the best suited to receive these refugees thanks to the the already significant Yugoslavian community residing there and to Germany's ability and willingness to provide protection to those fleeing the war.¹²⁰ The flow of refugees into Germany responded to the dynamics of the conflict: in the early stages of the war, involving mostly Croatians, about one hundred thousand of them arrived to Germany; later on, when the war spread to Bosnia, acts of systematic violence triggered massive outflows from those areas and Germany hosted some 350,000 Bosnian refugees. Simultaneously, Germany also received another 250,000 Yugoslavians mainly from Serbia and from Kosovo. Thus, overall during the first half of the 1990s, Germany received roughly 700,000 migrants from Yugoslavia, making it by far the largest recipient foreign country (see Lederer (1997) for a detailed account of these flows).

¹¹⁷The paper is accompanied by an Online Appendix, which is referred to throughout the text.

¹¹⁸From 1998 to 1999 the region was affected by yet another armed conflict: the Kosovo War. Our focus, however, is mostly on the conflicts that occurred prior to that and which resulted in massive population displacements.

¹¹⁹See Angrist and Kugler (2003) for a summary of migration of Yugoslavian nationals to different European destinations (in the context of a study on the impact they had on local labor markets).

¹²⁰Throughout the paper, we refer to all those people escaping the Yugoslavian civil war as "refugees". This is a much broader use than the legal definition of refugee, which implies having being recognized by a receiving country, on a case-by-case basis, as a refugee (i.e., having an asylum request approved) according to the definitions agreed upon and stated in the Geneva Convention on Refugees of 1951 and the Protocol Relating to the Status of Refugees of 1967 (among other country or region-specific definitions).

Yugoslavian refugees that arrived to Germany were given the option of acquiring a temporary protection status, known as *Duldung* –a status created by the German authorities at the time¹²¹– which can be translated to English as "toleration". The temporary character of the *Duldung* status did not constitute a permanent residence permit, but rather it was a "suspended deportation" status. In other words, a *Duldung* holder was allowed to remain in Germany until the *Duldung*'s expiration, after which its holder was obliged to leave the country immediately. While the *Duldung* duration upon issuance was set to six months, the authorities had the option to renew it for another six months period as long as it was not safe for the refugee to return home (Dimova, 2006). De facto, the *Duldung* status was renewed for all holders as long as the war was still going on. *Duldung* holders were allowed to work in Germany shortly following their arrival, and had full mobility rights within the country.

Another less popular option for Yugoslavians fleeing the war and arriving to Germany was to formalize their refugee status, or in other words, to request and then receive formal asylum. According to Article 16(a) of the German Basic Law (*Grundgesetz*), an individual is eligible to seek asylum if he or she faces individual persecution and is able prove so. If granted asylum, the individual enters a path towards permanent residency (Hailbronner, 2003). Asylum recognition rates, however, were very low for citizens from the former Yugoslavia.¹²² This is because most of them could not prove to the German authorities they were facing individual persecution at home following the standards set by the German authorities at the time (Dimova, 2006). Importantly enough, however, asylum seekers whose asylum request was denied were eligible to receive the *Duldung* status, and most of them did.

4.3 Labor Market Conditions and Mobility of Refugees

A large number of Yugoslavian refugees managed to integrate into the German labor force upon arrival.¹²³ Overall, both *Duldung* holders and asylum seekers (e.g., those waiting for their asylum application to be approved or denied) had free access to the local labor market shortly after their arrival, with some important differences. *Duldung* holders were entitled to receive a work permit, and were allowed to work in Germany without any geographical nor sectoral limit. Asylum seekers who arrived prior to 1997 were also allowed to work three months after their arrival date and while their application was being considered (if the application was eventually accepted, they naturally kept their right to work).¹²⁴

Yet, an important difference between the two statuses concerned the mobility of the workers: *Duldung*

¹²¹In the early 1990s, as a response to the legal difficulties faced by the hundred of thousands of refugees seeking protection, –and knowing that most of these people were not eligible for asylum– the German government created the *Duldung* status. *Duldung* was granted relatively quickly to all those arriving because of the war in Yugoslavia, making it possible for Germany to process large numbers of arrivals. Compared to other European countries this was a considerable humanitarian gesture on the side of the German government.

¹²²Between 1992 and 1995 only 1 percent of Bosnian applicants were granted asylum (Lederer, 1997).

¹²³The number of employed Yugoslavians rose by 60,000 people in 1992, equivalent to a 15.3% increase, as compared to 1991 (Deutscher-Bundestag, 1994).

¹²⁴Labor market access conditions for asylum seekers changed a few times. Until 1991, immediate access to labor market was possible. Between 1991-1997, a waiting period of three months was enacted. Modifications in the law in 1997 banned asylum seekers from the labor market. This changed in 2001 when 1-year waiting time was introduced. For more details see Liedtke (2002).

holders could live and work with no within-country geographical limitation whatsoever. Asylum seekers, on the other hand, were subjected to mandatory residency (*Residenzpflicht*) while their application was considered. They were obliged to stay within the region in which their application was processed.¹²⁵ The decision on which region would process the application was made by the authorities based on pre-established quotas. This limitation on geographic mobility for asylum seekers is an important part of our identification strategy, which we detail in Section 4.6.1.

4.4 End of the War and Deportation

The signature of the Dayton Peace Accord in December 1995 officially marked the end of the war that started in 1991 (in particular, by putting an end to the Bosnia war). After that date, the German authorities had no reason to further renew the *Duldung* status of refugees and indeed enacted the imminent deportation of refugees back to the former Yugoslavia.¹²⁶ In fact, only one day after the signing of the Dayton Accord, Germany formally announced a repatriation plan through which *Duldung* refugees were gradually forced to leave the country (Dimova, 2006), often simultaneously rolling out assisted repatriation programs (Bosswick, 2000).¹²⁷

Repatriation was planned in two main phases. The first phase targeted single adults and childless couples as well as people with family back in their home country. The second phase targeted the rest of the refugees. By the summer of 1996 letters requesting deportation were sent out, and by the end of 1996 people started getting deported. Repatriation and deportations continued until 2000, though most of them had happened by 1998. Figures by international organizations and independent academic research suggest that about 75 percent of Yugoslavians civil war refugees returned to their home country or to another former Yugoslavian nation, with an additional 15 percent settling in third countries and only about 10 percent remaining in Germany (UNHCR, 2005; Rühl and Lederer, 2001; Lederer, 1997).¹²⁸

Refugees found employment across diverse sectors and relied on different channels to secure their jobs. Some were able to utilize their network of friends and family relatives, some relied on local employment agencies and some found work by themselves (e.g., Walker, 2010; Ruben et al., 2009). Ruben et al. (2009)

¹²⁵The rules on mobility while the application was being processed were defined by local governments. Some states restricted movement of the asylum seeker to a district, while others allowed free mobility within the state.

¹²⁶For the Croatians, however the deportations started following the signature of the cease-fire agreement known as the Vance Plan in January of 1992. By the end of 1994 almost all of the Croatian refugees had returned (Lederer, 1997).

¹²⁷Voluntary returns were mainly realized as a part of the program of German Government through REAG (Program for Reintegration and Emigration for Claimants of Asylum in Germany) and GARP (Government Program of Assistance to Repatriation) which was implemented in cooperation with the International Organization for Migration (IOM) whose target was to support voluntary return. Both programs were completed in 2001 (Nenadic et al., 2005).

¹²⁸These numbers are confirmed when looking at the return rates with respect to the different nationalities and ethnicities involved in the conflict. For example, when it comes to Bosnian refugees, Rühl and Lederer (2001, p.50) describe: “[t]he number of Bosnian war refugees fell from 345,000 to approximately 28,000 by December 2000, more than 260,000 of which went voluntarily to Bosnia-Herzegovina. About 51,000 have migrated on to other countries (to the USA, Canada and Australia). The proportion of forced repatriations is well below 2% (approximately 5,500 cases).” With respect to Croatian refugees, Lederer (1997, p.310) explains: “During the Croatian-Serbian War (1991 to 1993) numerous Croatians were also admitted to the Federal Republic of Germany. According to information from the Federal Ministry of the Interior of 9 October 1996, most of the original 100,000 Croatian refugees should have returned to their homeland within the framework of the repatriation process that began in 1994. However, the Federal Ministry of the Interior notes that there is no precise information on this from the federal states.”

note that those who worked during their time in Germany were also more likely to be employed upon return to the former Yugoslavia, with many finding jobs in industries in which they had worked before the war, others being employed in related sectors and some in totally new activities. There is anecdotal evidence, however, suggesting that after returning to their home countries refugees subsequently worked (or founded companies) in the very same sector they had worked at in Germany (or other countries such as Sweden or Austria). In Online Appendix Section 4.10.1 we give four examples of anecdotal stories presenting such professional itineraries. As these stories show, refugees benefited from their experiences in Germany in many ways. In addition to acquiring knowledge and skills in new industries, they learned about different production methods, and established networks which they later put into use to create trade links and attract foreign investment when establishing their own companies. Once back home, most of them continued working in the same sectors where they had been employed during their stay in Germany. Although some found jobs as regular workers, many chose to set up their own companies.

4.5 Data and Sample

We link a number of datasets together for our study. First, we use data on exports for the period 1984-2014 which comes from bilateral trade data compiled by Feenstra et al. (2005) with extensions and corrections suggested by Hausmann et al. (2014) (the original source of the trade data is UN Comtrade). In most cases our dependent variable is exports by product from each country to the rest of the world excluding Germany. We do this so that our results are not confounded with an increase in trade driven by lower transaction costs caused by migrant networks (e.g., Rauch and Trindade, 2002; Parsons and Vézina, 2018a).

Products are defined using the 4-digit Standard Industry Trade Classification (SITC) revision 2, and include 786 different varieties. This product classification provides a disaggregation level that enables a meaningful discussion about export diversification patterns. Some examples of products in this level of disaggregation are, for example, "Knitted/Crocheted Fabrics Elastic or Rubberized" (SITC 6553), or "Electrical Measuring, Checking, Analyzing Instruments" (SITC 8748). Following Hausmann et al. (2014), we exclude countries below 1 million citizens and total trade below USD \$1 billion in 2010. Other variables created using trade data are explained as they are introduced into the analysis.

The data on migrant workers in Germany are based on records from the German social security system and comprise all persons employed subject to social security contributions, with the exception of self-employed and civil servants.¹²⁹ The records indicate the industry where the workers are employed. Our sample is restricted to 40% random draws of foreign nationals observed on June 30 of each year from 1975 to 2014 augmented by the employment history of each individual for our sampling period.

¹²⁹These records have been assembled by the Institute of Employment Research (IAB) into the Employment History (BeH) data file (IAB employment history (BeH), 2015). The data or variants of it have been widely used to study a variety of labor market aspects (e.g., Card et al., 2013; Dustmann and Glitz, 2015; Dustmann et al., 2017b).

This amounts to about 2.4 million workers per year on average, which is a large enough amount for the random draws to form a representative sample.¹³⁰ Moreover, since we can observe the full employment history, we can determine whether an individual was employed before or after any given year in Germany, which we exploit to construct our treatment. Beyond individual information such as age, nationality, and educational attainment, the data include detailed occupational codes categorized in more than 300 different occupations.¹³¹

We also collected information on direct investments of German firms in former Yugoslavia sent to us by the German Central Bank (*Deutsche Bundesbank*) upon request. We use these data to compute German FDI stocks in former Yugoslavian countries at 2-digit SITC level between the years 1990-2010, to serve as a control. Second, in order to construct an instrument to deal with endogeneity issues, we used data on asylum applications in Germany, which comes from the German Federal Office for Migration and Refugees (*Bundesamt für Migration und Flüchtlinge, BAMF*) sent to us upon request, too. We also use inflow quotas mandated by the government that define the regional distribution of asylum seekers (*Königsteiner Schlüssel*). These quotas are determined yearly by the Joint Science Conference (*Gemeinsame Wissenschaftskonferenz, GWK*). The yearly data between 1990 and 2016 have been sent to us by the GWK upon request.

With these datasets we are able to match the exports of Yugoslavia to the rest of the world with the number of Yugoslavian workers in Germany working in the same product category.¹³²

4.6 Natural Experiment: Yugoslavian Refugees

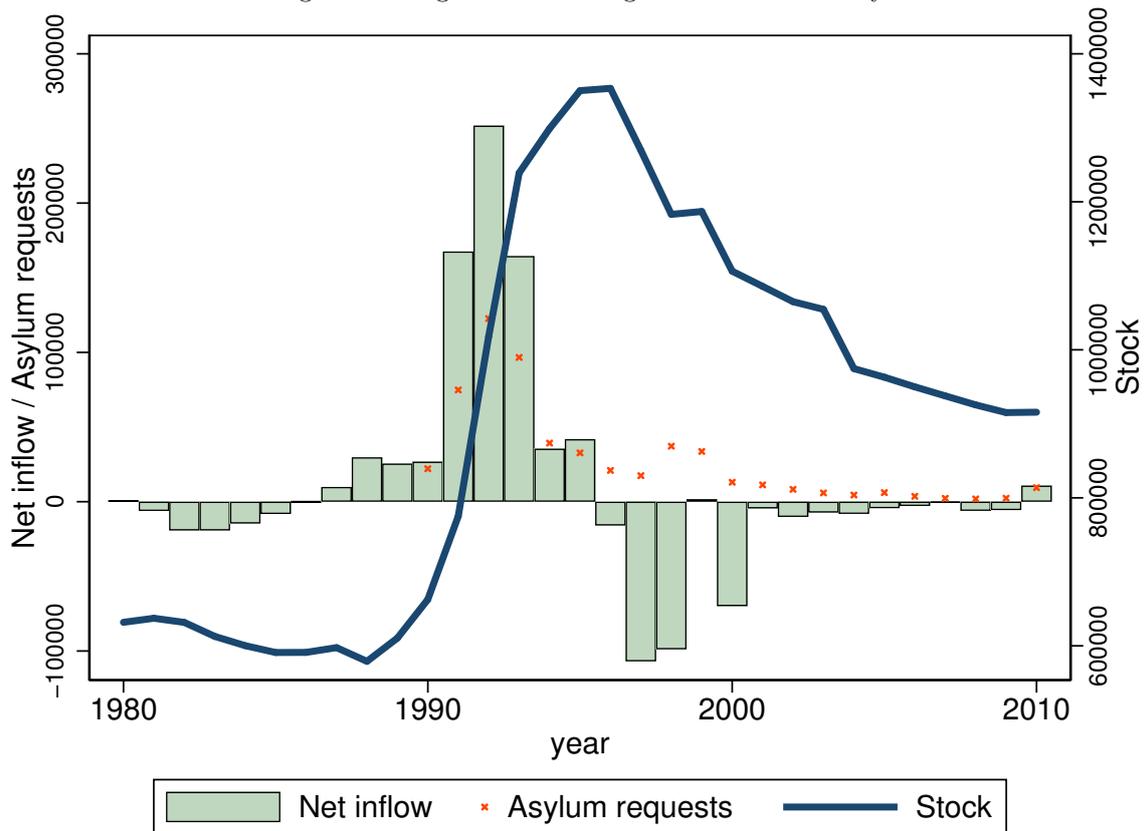
As pointed out in Section 5.3, around 700,000 migrants from the former Yugoslavia arrived to Germany in the first half of the 1990s. Figure (24) summarizes these numbers. In 1980 there were already about 600,000 Yugoslavians residing in Germany. This stock remained steady until the late 1980s when the net inflow of Yugoslavian migrants started to grow at a rate of 25,000 per year, including until the year 1990. This rate skyrocketed to 168,000, 250,000 and 165,000 during 1991, 1992 and 1993, respectively. The sharp increase in the net inflow of migrants was fueled by refugees escaping war. We also see a sharp increase in asylum requests from Yugoslavian citizens during the same years.

¹³⁰For privacy issues, the sample utilized in this paper is an anonymized version of the original database. In order to comply with data privacy rules, sensitive values –industry-period observations with less than 20 workers– have been replaced with different moments of the distribution of the number of migrant workers in the same industry and year. The number of cells affected depends of course on the level of disaggregation of worker characteristics such as education, occupation, skill etc. More details on this procedure can be provided upon request. The results presented herein, however, are robust to using the non-anonymized version instead.

¹³¹See more details on this dataset in Online Appendix Section 4.10.2.

¹³²Using our employment sample we compute the number of workers in Germany by nationality and year for all SITC 4-digit product categories. We use the nationality of the worker based on the passport recorded at his or her first appearance in the database. To compute the number of workers by product we rely on the concordance tables produced by the United Nations Statistical Office and the work by Dauth et al. (2014) that matches German 3-digit WZ industry codes to 4-digit SITC (rev. 2) products. When match is one-to-many, we create our own concordance following the same procedure as Cuñat and Melitz (2012) described in their footnote 24. In particular, we use the distribution of German exports in 1995 as a proxy for the distribution of employment across different 4-digit SITC products that correspond to a single 3-digit WZ German industry. Further details on the employment sample, variable descriptions and auxiliary data are provided in Online Appendix Section 4.10.2.

Figure 24: Migration from Yugoslavia into Germany



The figure shows the net inflow, stock and asylum requests of migrants from (former) Yugoslavia into Germany, from 1980 until 2010. The number of migrant stocks by nationality is based on the Central Register of Foreigners (*Ausländerzentralregister, AZR*). The data have been downloaded from the GENESIS-online data base of the German Federal Statistical Office (*Statistisches Bundesamt*), Table 12521-0002. Data on migration flows by nationality are from the migrations statistics (*Wanderungsstatistik*) of the German Federal Statistical Office (*Statistisches Bundesamt*) and sent to us upon request.

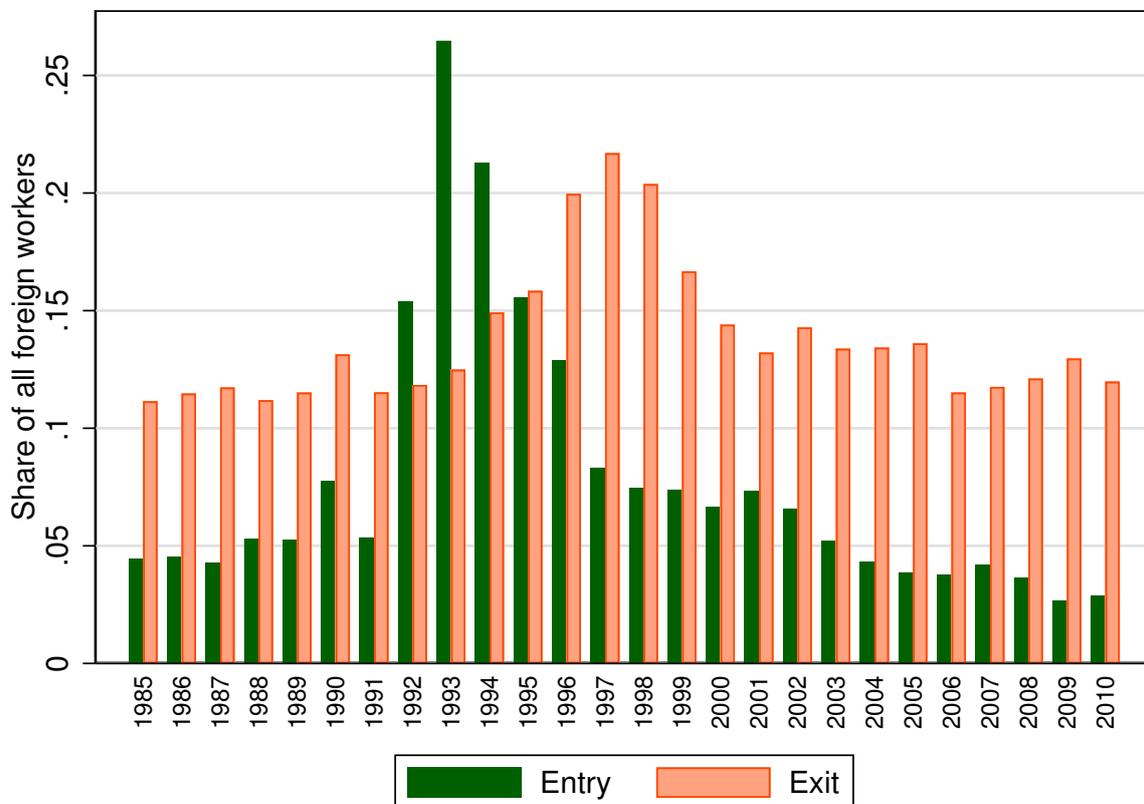
The number of Yugoslavians in Germany sharply declines starting in 1996, after the Dayton treaty was signed. By 2000 close to 350 thousand Yugoslavians had left the country. While some of them left to a third country, it has been estimated that the majority of them returned to countries of the (by then) former Yugoslavia (UNHCR, 2005; Ruhl and Lederer, 2001; Lederer, 1997).

The natural experiment, however, uses data on Yugoslavian workers that joined the German labor force (in each tradable sector) between 1991 and 1995 and had dropped out of it by year 2000. These parameters allow us to identify, with high probability, those Yugoslavians of working age who were *Duldung* holders and had been deported by year 2000.¹³³ In our data we see patterns consistent with the historical narrative described so far: both the inflow of Yugoslavian workers into the tradable sector

¹³³Our data also show that about half of these Yugoslavian workers who arrived between 1991 and 1995 are still active in the German labor force by 2014. Presumably, these "stayers" were not *Duldung* holders (and therefore, were not subject to deportation) or, alternatively, they were *Duldung* holders but were allowed to stay for humanitarian reasons. Again, as we noted above, the overall statistics suggest that over 75 percent of *Duldung* holders returned to the newly established countries of the former Yugoslavia. These are the people our treatment aims to capture: our treatment counts those workers who arrive between 1991 and 1995 and drop out of the labor force by the year 2000, by industry. This means that we are in fact counting all those workers in each manufacturing industry that were, most likely, *Duldung* holders. The fact that some other Yugoslavians who stayed beyond 2000 for whatever reason is not a threat to our identification strategy, as long as these cases were not more or less frequent in some industries than in others. We expand on this in Section 4.6.1.

labor force between the years 1991 and 1995 and their outflow by year 2000 is highly unusual as compared to foreign workers from other nationalities, as Figure 26 shows. The figure graphs the yearly share of Yugoslavian workers entering and exiting the labor force of Germany's tradable sector. The share is always computed using the total number of all foreign workers entering and exiting the labor force in those same years at the denominator. It becomes clear from looking at the figure that Yugoslavians entered the labor force in much higher proportion during the first half of the 1990s, as compared to the same proportion in years before 1990 and after 2000. We also see that Yugoslavian workers exited the German labor force in higher proportion during the later half of the 1990s, consistently with the historical events.

Figure 25: Yugoslavian workers yearly entry to and exit from German's labor force



The graph shows the yearly share (out of all foreign workers) of Yugoslavians entering and exiting the labor force of Germany's tradable sector.

4.6.1 Empirical Strategy

Our objective is to study changes in product-level Yugoslavian exports to the rest of the world given different levels of return migration of Yugoslavian refugees who worked in Germany in the corresponding industry. We do this through a difference-in-differences estimation. Given that the German data do not allow us to distinguish which is the region of origin of the refugees within Yugoslavia (we only see that they entered the labor force with a Yugoslavian passport), our unit of analysis is the combined exports

by product of all countries of the former Yugoslavia. That is, the trade data includes export by product of Yugoslavia as a nation until 1991, and we complement this by simply adding up exports by product of all countries that formed Yugoslavia post 1992: Bosnia and Herzegovina, Croatia, FYR of Macedonia, Montenegro, Serbia and Slovenia.¹³⁴ We end up having a balanced-panel of exports by product for the former Yugoslavia from 1984 until 2014, which is the main input required to construct our dependent variable.

The main independent variable –the treatment– is the number of Yugoslavian refugees by sector who (i) joined the German labor force between 1991 and 1995, (ii) had not been recorded in our data in 1990 or before, and (iii) had not been recorded in our data in 2000 or after.¹³⁵ We cannot distinguish whether these workers with Yugoslavian passports that left the labor force indeed returned back to the former Yugoslavia. Thus, in our calculation of return migration we are including workers who, for instance, stayed in Germany working in the informal sector or went to a third country. Yet, all these possibilities work against us in our estimation, implying that our estimates are to some extent understating the true effect of return migration. Moreover, the historical context summarized above suggests that about 75 percent of those who were repatriated upon the expiration of their *Duldung* status returned to the former Yugoslavia.¹³⁶

Figure 26 describes the treatment variable. It plots the number of Yugoslavian workers that entered the German workforce between the years 1991 and 1995 (horizontal axis), and the number of those workers who remain in the labor force beyond year 2000 (vertical axis), by 4-digit SITC code. All observations are below the 45 degree line, simply because the number of migrants who stay in each industry beyond the year 2000 is a subset of all those who arrived between 1991 and 1995. Thus, the treatment is the difference between each one of the observations and the 45 degree line. As can be seen in the graph, there is variation in the treatment across industries. Some of the codes that stand out as having a large amount of workers returnees are 8219 (Furniture parts), 6911 (iron and steel structures), 5989 (chemical products), and 2482 (worked wood of coniferous). In our data organized by industries we notice that, among those who were employed at some point as wage earners in the manufacturing sector, return rates were substantial but not as high for the average refugee as documented in official reports from international organizations (e.g., UNHCR, 2005). On average, about 30 percent of the workers that arrived between 1991 and 1995 had dropped out of the labor force by 2000.¹³⁷

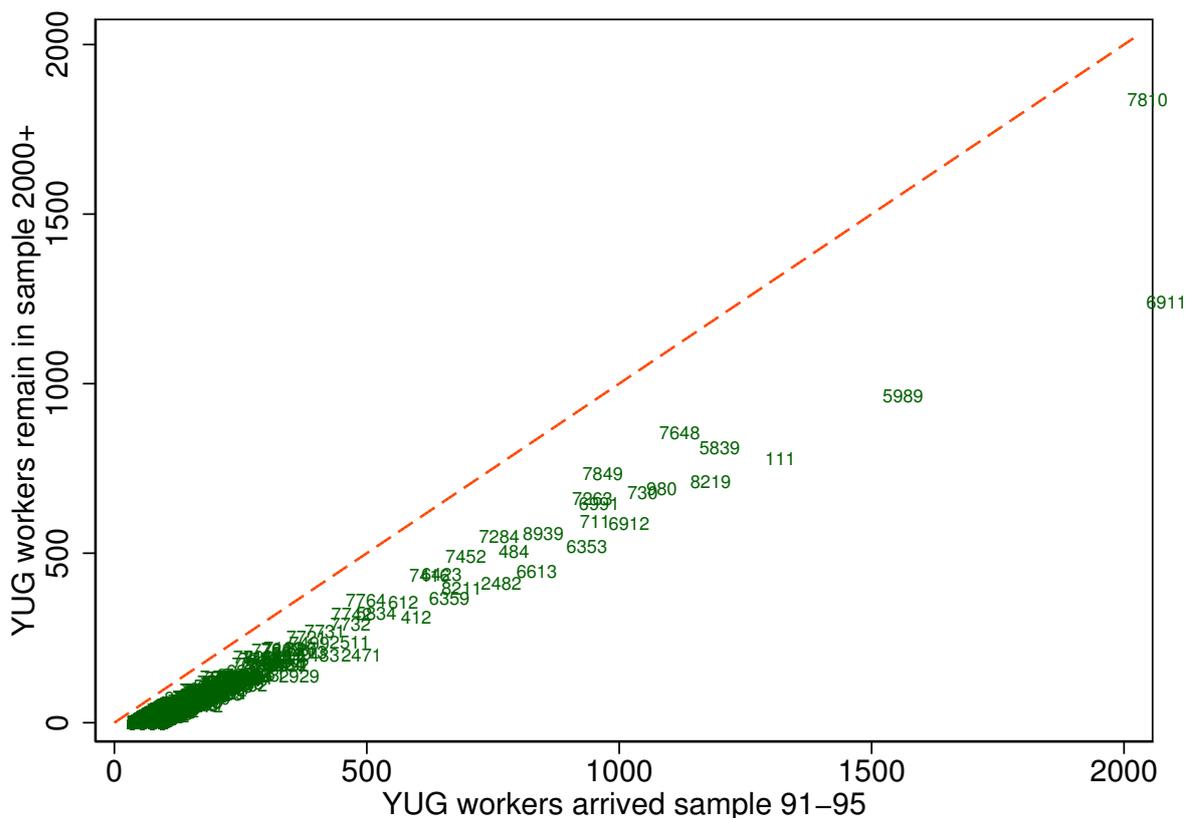
¹³⁴Very few people left Slovenia while almost none left the FYR of Macedonia as both countries obtained their independence with limited or no armed conflict. While Slovenia was the republic with highest GDP per capita and a much more diversified export basket than the rest of the countries to begin with, Macedonia was one of the poorest republics of the former Yugoslavia with little exports. Our results are robust to excluding both Slovenia and Macedonia from the exports data (see Online Appendix Section 4.10.9).

¹³⁵Finding no entry for a person in our data implies that this person was not employed in any job, industry, or occupation subject to social security contributions on June 30 of any given year.

¹³⁶About 10% of the Yugoslavian workers we see entering the labor force between 1991 and 1995 were 18 years or younger at the year of entry. 75% of them, in contrast, were 20 or older and 50% were 25 or older. This rules out the possibility that the entry of Yugoslavian into the labor force is mostly driven by locals with Yugoslavian passports joining the labor force at a young age, rather than by refugees arriving to Germany.

¹³⁷However, when focusing on those migrants that we can identify in our data as Bosnians, even though they are a small

Figure 26: Yugoslavians in German workforce, by product

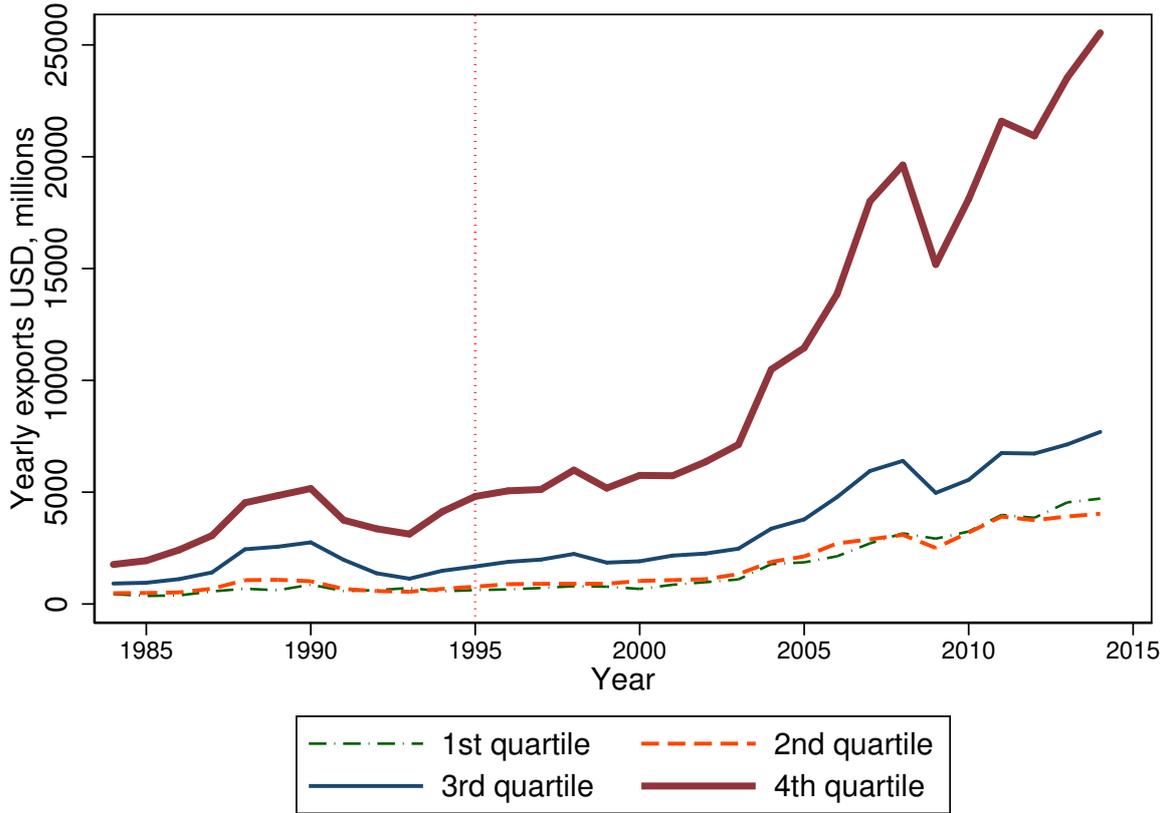


The figure shows the number of Yugoslavian workers in the German workforce that arrived between 1991 and 1995 against those that remain in year 2000 and beyond, by 4-digit product.

Before we turn to the econometrics, we look at whether products associated with a larger reduction in the number of Yugoslavian workers in Germany experienced better export performance in the former Yugoslavian countries upon their return. Using only raw data, Figure 27 visualizes the total value of exports of products linked to different levels of treatment (i.e., quartiles), year after year. The figure shows that up until 1995 (the year where our "treatment" begins) products in the four different quartiles had somewhat parallel trends. However, after 1995, the third and fourth quartiles in terms of treatment intensity diverge quite significantly from the first two quartiles. This visualization not only provides some descriptive evidence of the results holding with raw data, but also shows that the parallel trends assumption required for the difference-in-differences methodology is a reasonable one. In any event, we address pre-trend issues more thoroughly in the next section.

share, the drop-out rate of the labor force is more than 50 percent, closer to the global official figures.

Figure 27: Exports for products with different levels of treatment



The figure plots the cumulative value of exports of the former Yugoslavia to the rest of the world (vertical axis) across years. Treatment is defined as the number of return migrants from Germany by 2000.

We now turn to regression analysis and estimate the following difference-in-differences specification:

$$exports_{p,t} = \beta^{DID} treat_p \times after_t + \beta^C DEUexpshare1990_p \times after_t + \eta_p + \alpha_t + \varepsilon_{p,t} \quad (36)$$

Subscripts p and t represent product and year, respectively. The left-hand side variable ($exports_{p,t}$) measures the value of exports from the former Yugoslavia to the rest of the world excluding Germany for product p during year t .

We start by estimating this regression using two periods: "before", which corresponds to 1990, just before the war started, and "after", which corresponds to 2005, five years after most Yugoslavian refugees had returned. To avoid our results being driven by noise in the dependent variable for a particular year, we use average exports per product between 1988 and 1990 for the "before" period, and the average between 2005 to 2007 for the "after" period.¹³⁸

The variable of interest $treat_p$ is the number of workers that joined the German labor force for product p between 1991 and 1995 and had dropped out of that same labor force by 2000, according to our definition earlier. We rescale the variable $treat_p$ using the inverse hyperbolic sine. The inverse

¹³⁸Our results, however, are robust to using only data for the actual years for which the "before" and "after" periods are defined: 1990 and 2005.

hyperbolic sine is defined at zero and behaves similarly to a log-transformation. The interpretation of regression estimators in the form of the inverse hyperbolic sine is similar to the interpretation of a log-transformed variable (see MacKinnon and Magee, 1990).¹³⁹ $after_t$ is a dummy variable which equals 1 for the observations corresponding to the "after" period.

Given that in the actual estimations both the dependent and the main independent variables are in a logarithmic-type form (details on this below), β^{DID} represents the elasticity of exports to returnee workers. That is, industries with a 1 percent larger pool of returnee workers have larger export value by β^{DID} percent larger between the "before" and "after" periods compared to industries with no returnee workers. Bear in mind that since this is a difference-in-difference setting, our results reflect relative differences in levels across industries based on their exposure to the treatment.

We often include controls to avoid our results being driven by other factors not accounted for. One of those controls which forms part of our main specification is an interaction term between the relative size of each product p in Germany's export basket in the "before" period (1990) and the "after" binary term. This interaction allows to control for differential trends based on the possibility that the export basket of Yugoslavian countries in 2005 evolves towards the export basket of Germany in 1990, due to structural transformation convergence processes. If convergence is something to be worried about in terms of that process driving our results, and not the migrants, then this control will take care of that possibility.

As for the other terms: η_p represents product fixed effects while α_t represents year fixed effects (which in the main estimation is equivalent to one dummy variable for the year 2005). The two fixed effects are perfectly multi collinear with the terms $treat_p$ and $after_t$, and so there is no need to add the interacted terms separately. $\varepsilon_{p,t}$ represents the error term. Our estimations cluster standard errors at the product level, the level at which the treatment varies (Besley and Burgess, 2004; Bertrand et al., 2004).

Using this specification, claiming that our results are unbiased imply that the industry-specific entry to and exit from the German labor force of Yugoslavian workers are exogenous with respect to present and future dynamics of Yugoslavian exports. Can we say this is indeed the case?

Our identification relies on the exogeneity of arrival and exit of refugees into and out of the German labor force with regards to export trends back in Yugoslavia at the industry level. There are two main endogeneity concerns in our specification. First, the possibility that upon arrival, refugees self-selected into particular tradable sectors with growth potential in Yugoslavia based. Second, the possibility that the exit of refugees from the German labor force by year 2000 –even if it was enforced by across-the-board repatriation efforts by the German authorities– happened more or less frequently in particular industries in a way that is correlated with export dynamics in Yugoslavia. We address each of these concerns below.

¹³⁹The inverse hyperbolic sine ($asinh$) is defined as $\log(y_i + \sqrt{(y_i^2 + 1)})$. Except for small values of y , $asinh(y_i) = \log(2) + \log(y_i)$.

Self-selection into industries upon arrival

In order to deal with the possibility that migrants self-selected into particular industries in a way that correlates with future Yugoslavian exports, we construct an instrumental variable that computes expected worker stocks per industry by exploiting a spatial dispersal policy applied to Yugoslavian asylum seekers

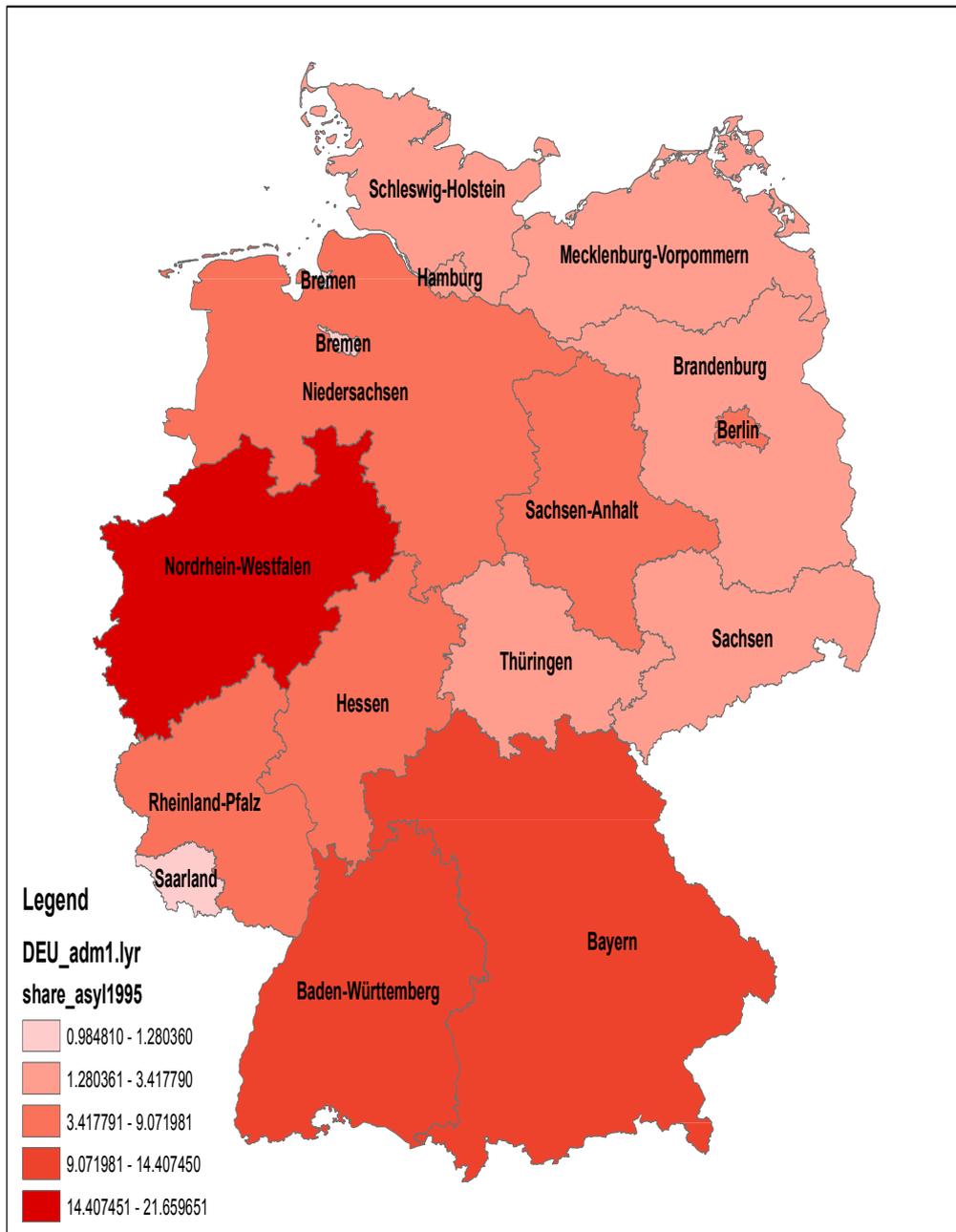
While asylum requests were being processed, asylum seekers were sent to different parts of the country following the Königstein State Convention (*Königsteiner Staatsabkommen*) which was signed in 1949 by all German federal states and defined cost-sharing rules between states in jointly financed projects. Although initially this convention concerned financing of joint science projects, the system was later adopted –among other things– for the allocation of asylum seekers within States in Germany. The dispersal of asylum seekers is regulated at the federal level by the Asylum Procedure Act (*Asylverfahrensgesetz*), where each state is allocated a certain number of asylum seekers according to its “Königstein” quota (*Königsteiner Schlüssel*). The quota is based on the weighted sum of population (1/3) and tax revenues (2/3), and it is re-calculated annually. In the absence of substantial regional shocks, this quota does not vary much over time.

In particular, upon the arrival of an asylum seeker into the German territory, he or she is absorbed by a reception center in the Federal State of arrival if there is any remaining capacity to receive more people according to the quota described above, or alternatively, the person is allocated to the reception center in a Federal State with the most vacancies according to the quota.¹⁴⁰ Our identification strategy is based upon the premise that this allocation was exogenous.

An illustration of the quota system can be seen in Figure 28, which shows the share of asylum seekers that each of the sixteen German states should have received using the quota system in year 1995. For example, Nordrhein-Westfalen is the state that should have received most of the asylum seekers in 1995, followed by Baden-Württemberg and Bayern, while states such as Bremen or Saarland received a very small share.

¹⁴⁰If more than one reception center fits these criteria, the geographically nearest one to the entrance location of the asylum seeker is appointed responsible. After the first period in reception facilities, which can last up to a maximum of six months, the asylum seeker is placed in a district within the state of first allocation. The residence obligation ends as soon as the Federal Office grants asylum status. The average duration of the application procedure was between six months and two years.

Figure 28: Distribution rule of asylum seekers in Germany 1995



The figure maps the different German states with their shade representing the share of all asylum seekers in Germany they were mandated to receive by law in 1995, based on their population and tax revenues.

Thus, to construct the instrument we combine three pieces of data: (1) the yearly inflow of asylum seekers from Yugoslavia to Germany between 1991 to 1995, (2) the yearly asylum quotas for German states,¹⁴¹ and (3) the relative size in terms of employment in each industry in each state of Yugoslavian workers in 1975, with the intuition that the distribution of Yugoslavian migrants in 1975 can play a

¹⁴¹In fact, since we don't have data pre-1995 for states of eastern Germany, we set those at zero for years 1991 to 1994. However, this lack of variation is not critical, turns out. According to the employment data, in 1995 there were over 367 thousand Yugoslavians employed in western German states across all industries, as compared to only 1400 in eastern Germany, or just 0.38%. Thus, that lack of variation should not affect the relevance of our instrument.

role in explaining the industry-level integration labor market integration of the asylum seekers in the early 1990s.¹⁴² The resulting variable estimates the number of Yugoslavian asylum seeking workers in Germany per industry. The following equation reflects the calculation:

$$\underbrace{TreatIV_p}_{\substack{\text{Expected asylum} \\ \text{seekers workers in } p}} = \sum_{t=1991}^{1995} \sum_s \underbrace{asylumseekers_t}_{\substack{\text{Inflow asylum} \\ \text{seekers in year } t}} \times \underbrace{quota_{s,t}}_{\substack{\text{Quota (share) for} \\ \text{state } s \text{ during } t}} \times \underbrace{YUGshareindustry_{s,p,1975}}_{\substack{\text{Yugoslavians employment share} \\ \text{of } p \text{ within } s \text{ in 1975}}}$$

The instrumental variable is a feasible one under two conditions: first, if it correlates with the treatment, and second, if the exclusion restriction holds. In terms of the first condition, we see a strong correlation between the treatment and the instrumental variable. Figure 29 shows plots $TreatIV_p$ in the horizontal axis against the number of Yugoslavian workers that had left the labor force between 1995 and 2000 (our treatment). Each observation in the figure is a product, symbolized by its 4-digit SITC code. It can be seen in the figure that the number of Yugoslavian workers who left the labor force between 1995 and 2000 in each industry are strongly correlated with stocks that are predicted through the spatial dispersal policy.

Figure 29: Instrumental Variable Relevance



The figure plots the expected number of asylum seekers expected to work in each industry based on their geographic allocation in each state and the employment share of each 4-digit SITC code in that state using data from 1991 to 1995 in the horizontal axis (in logs) against the number of Yugoslavian workers who arrived between 1991 and 1995 and leave the German labor force between by 2000 (in logs), by each 4-digit SITC code. The figure represents a graphical visualization of the first stage of the two stages least squares.

There are many reasons for which this strong correlation is not surprising. Given migrant networks,

¹⁴²We are aware of the critique by Jaeger et al. (2018) regarding using past spatial distributions of migrants to instrument for current distribution, though in our paper it lacks relevance given that our dependent variable does not reflect economic activity in the same location of the migrants' destination but rather in their country of origin.

we can expect the spatial distribution of Yugoslavian workers across states and industries in 1975 explains industry-level integration of asylum seekers during the early 1990s. Limiting the instrument to asylum seekers (due to their exogenous spatial allocation) –even if they were a small share of all refugees as seen in Figure 24– can explain the overall industry-level distribution of all Yugoslavian refugee workers responds to two facts. First, the geographic allocation of asylum seekers is relevant for all refugees who request for asylum, even if the asylum turns out not to be approved. That is, all refugees who eventually got a *Duldung* status but who originally requested for asylum, had to comply with this exogenous geographic allocation while their asylum status was being reviewed by the authorities. Second, the exogenous allocation of the share of Yugoslavians who actually requested asylum might as well be explanatory of the location choice of those who received *Duldung* even if they did not request for asylum to begin with, due to pull factors induced by pre-existing migrant networks.

Our main assumption regarding the second condition –the exclusion restriction– is that both the number of asylum seekers fleeing war and the quota of asylum seekers per state and year defined by the German federal authorities, as well as the relative size of the Yugoslavian workforce in each industry and state in 1975 are not correlated with post-1995 product-specific export trends of former Yugoslavian countries to the rest of the world other than through the migrants themselves. We have no reason to think that this assumption could be violated.

Exogeneity of exit with respect to industries

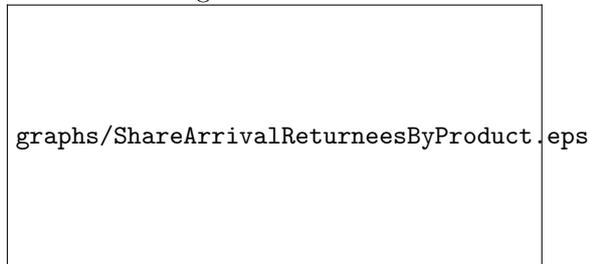
The other remaining concern is whether the exit of Yugoslavian *Duldung* holders from the German labor force was exogenous to the industry they were working in. For example, if workers left the labor force more massively in some industries more than others, in a way that is correlated to the future export potential in Yugoslavia, then this could invalidate our identification strategy.

This was not the case, as shown in Figure 30. The figure compares the proportion of Yugoslavians who arrived between 1991 and 1995 working in each 4-digit SITC product on the vertical axis, against the proportion of Yugoslavian who returned by 2000 (based on the definition of the treatment) working in each product on the horizontal axis. The dashed line represents the 45 degree line. If exit from the labor force by Yugoslavians was completely random, we would see a perfect alignment of those dots along the 45 degree line: the share of workers arriving into each industry must be the same as the share of workers leaving that industry. Barring some exceptions, the graph does approximate this scenario. In fact, the correlation between both shares is close to 0.9.

Most dots, each one representing a 4-digit SITC code, are quite consistently aligned along the 45 degree line. Take the industry code 6991, for example, which represents the sector "Locksmith hardware". About 17% of all Yugoslavian workers who arrived between 1991 and 1995 worked in that industry, according to the horizontal axis, and about 15% of those workers who left by the year 2000 dropped

from that product's labor force. All in all, the evidence suggests that the exit from the labor force was exogenous across industries. Figure 30 shows that there are a number of outliers (i.e., observations far from the 45 degree line) but these do not affect our results (see Online Appendix 4.10.6).

Figure 30: Share of arrival vs exit from the German labor force by product



The figure plots the share of Yugoslavian workers in each industry out of the total that joined the labor force between 1991 to 1995 on the vertical axis against the share of Yugoslavian workers in each industry out of the total that dropped out of the labor force by 2000 (according to the way we define the treatment).

4.6.2 Summary Statistics

Table 31 presents the summary statistics. Our sample includes 786 products, and since we use two points in time for the differences-in-differences estimation, the initial empirical analysis will use up to 1572 observations. The table presents summary statistics for the main variables in the regression. The first four lines of the table present data for the average export value from former Yugoslavian countries to the rest of the world in years 1990 and 2005, all in million US dollars (note that we don't adjust these values for inflation, which is accounted for by our year fixed-effect). These two points in time are the ones used in the main specification, which correspond to years before and after the war. However, we also present results for a multi-period analysis as well that uses export data for all the years in between.

Given the fact that the left hand side is calculated in US dollars, we are required to use a monotonic transformation to deal with the fat-tailed distribution. All of our results are presented using three different transformations: $\log(exports_{p,t})$, $\log(exports_{p,t} + 1)$ and $asinh(exports_{p,t})$. The first one is undefined for values where $exports_{p,t} = 0$, and therefore, when using $\log(exports_{p,t})$ as the dependent variable the sample size is reduced. The two other transformations deal with the occasions where $exports_{p,t} = 0$ by either adding USD \$1 before the transformation or by computing instead the inverse hyperbolic sine ($asinh$), respectively.¹⁴³

¹⁴³Since exports are aggregated across all destinations, the number of "zeroes" in the data is not as large as when using bilateral trade data. We explore this in detail in Online Appendix Section 4.10.3.

Table 31: Summary Statistics Yugoslavian Refugees in Germany

Variable	N	Mean	sd	Min	Max
Exports YUG in 1990, million USD	786	12.472	31.65	0.0	395.0
Exports YUG in 2005, million USD	786	24.458	71.62	0.0	1,090.0
YUG workers in 91-95	786	74.025	190.78	0.0	2,018.7
YUG workers in 91-95 & out by 2000	786	21.641	60.61	0.0	778.5

This table presents the sample summary statistics for the variables used to estimate specification (36).

Table 31 also summarizes the treatment. The third row in the Table presents statistics for the number of sampled workers with Yugoslavian passport that joined the German labor force at some point between 1991 and 1995. The average industry had 74 Yugoslavian workers that, arguably, arrived to Germany because of the war and joined the labor force. The next row is a subset of that group, and corresponds to our main treatment variable: the number of workers with Yugoslavian passport which had joined the German labor force sometime between 1991 and 1995 and had dropped from it by the year 2000. The value for this variable, averaged across all products, is 21.6. Our treatment exploits variation across industries, which we see in the table varies from 0 to 778.¹⁴⁴ All in all, our treatment is based on roughly 17,000 Yugoslavian workers across all industries, representative of the actual distribution.¹⁴⁵

Note that as mentioned above, our sample of working-age Yugoslavians employed in the tradable industry shows that the rate of return was roughly 30 percent, substantially lower than the anecdotal 75 percent figure, which applies to the Yugoslavian refugee population as a whole (UNHCR, 2005; Ruhl and Lederer, 2001; Lederer, 1997). This discrepancy, however, poses no problem for our identification strategy as long as the rate of return is not biased towards certain industries, which we discussed above in Section 4.6.1.

Also note that despite presenting the summary statistics in nominal values, unless otherwise stated, all right hand side variables are rescaled using the inverse hyperbolic sine for estimation purposes.

4.7 Results

Results for specification (36) are presented in Table 32. For the “before” period we use exports data averaged over 1988 to 1990, and for the “after” period we use exports data averaged over 2005 to 2007. The treatment is defined as the number of workers of Yugoslavian origin that left the German labor force between 1995 and 2000, by product.¹⁴⁶ The estimation includes product fixed effects, such that the results use only within-product variation, and year fixed effects, which in this case is equivalent to

¹⁴⁴Non-integer number of workers in an industry is a result of the use of weights based on industry code concordances during the data construction stage. For more information see Online Appendix Section 4.10.2.

¹⁴⁵210,000 Yugoslavian workers appear the first time in our data between 1991 and 1995. If the total flow was of 700 thousand people, it is reasonable that somewhere between a one-quarter and one-half of them were of working age. Our sample is, of course, smaller than the total population, thus the 210,000 figure seems reasonable. Of the 210,000 in our sample, 35% (or 75,000 workers) had exited the sample before the year 2000. Of those roughly 75,000, only 22% (around 17,000 workers) had a job in tradable industries during the 1990s.

¹⁴⁶Online Appendix Table 47 replicates the results using different treatments: return migration between 1995 and 2005 and the stock of migrants in 1995. The results are robust to using these different treatments.

a dummy variable for the year 2005. The first three columns report results using an OLS estimation, while the last three columns report results using a 2SLS estimation, making use of the instrumental variable described in Section 4.6.1. The table reports results using $\log(exports_{p,t})$, $\log(exports_{p,t} + 1)$ and $asinh(exports_{p,t})$ as dependent variables. Since the regressor $treat_p$ is rescaled using the inverse hyperbolic sine transformation –which behaves similarly to a log transformation–we interpret β^{DID} as an elasticity.¹⁴⁷

Table 32: Difference-in-difference estimation

Dependent variable: $exports_{p,t}$	OLS			2SLS		
	lnexp	lnexpplus1	asinhexp	lnexp	lnexpplus1	asinhexp
treat2000 \times after2005	0.0837 (0.038)**	0.1358 (0.063)**	0.1376 (0.066)**	0.1133 (0.051)**	0.2395 (0.086)***	0.2449 (0.089)***
germanexpshare1990 \times after2005	6.3680 (4.325)	-0.1571 (6.815)	-0.4415 (7.001)	3.3411 (5.054)	-11.3455 (12.364)	-12.0176 (12.800)
N	1496	1572	1572	1496	1572	1572
r2	0.86	0.81	0.81	0.86	0.81	0.81
KP F Stat				661.95	723.55	723.55

This table shows result of the estimation for specification (36) using different monotonic transformations for $exports_{p,t}$ in each column. The estimation uses average exports for years 1988 to 1990 in the "before" period and average exports for years 2005 to 2007 in the "after" period. The first three columns report results from an OLS estimation, while the last three columns report results from a 2SLS estimation. All columns include product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In the first three columns, we find all estimates to be positive and statistically different from zero for all different monotonic transformations of the dependent variable. The standard errors are clustered at the product level, which is the level of disaggregation of the treatment.

Column 1 of Table 32 presents the estimates when using the natural logarithmic transformation for the dependent variable. The point estimate in the first column is around half the size of those in the other two columns. This is not surprising as the first column excludes zeros and therefore excludes instances in which products are more likely to grow faster if they have a non-zero value in the second period.¹⁴⁸ Yet, this difference says something more: the fact that the point estimates in columns 2 and 3 are positive and significant –which include instances where a product was inexistent in the export basket of Yugoslavia by 1995–, and are larger than the point estimate in column 1, implies that the effect of return migration on comparative advantage is valid at the extensive margin (e.g., opening a new line

¹⁴⁷The continuous character of our treatment implies, arguably, that our estimator can be characterized as a fuzzy differences-in-differences one (see De Chaisemartin and D'Haultfoeuille, 2018). In our setting, the "control" group is stable over time (e.g., there are no control group "switchers"), which implies our estimation only relies on the common trends assumption. In other words, our setting allows us not to require the "stable treatment over time", nor the "homogenous treatment effect between groups" assumptions (assumptions 5 and 6 in De Chaisemartin and D'Haultfoeuille (2018)). While relaxing assumption 5 in our setting is straightforward, doing the same with assumption 6 might not be. Thus, as a robustness test, we compute the Wald DID estimator following De Chaisemartin and D'Haultfoeuille (2018), defining the treated units those above the 25th percentile in terms of the treatment. We find our results reassuring: the Wald DID point estimates are between 0.15 to 0.28, depending on the monotonic transformation used, all statistically significant at the 10% level. The point estimates are slightly larger than the OLS ones reported in Table 32, but they all fall within the statistical margin of error of the estimators. We thank Clement De Chaisemartin for his guidance on this exercise.

¹⁴⁸In fact, Table 48 in the Online Appendix re-estimates columns (2), (3), (5) and (6) of Table 32 excluding observations with zero exports. In that case, the estimates are exactly the same as in columns (1) and (4) of Table 32.

of exports) as well as at the intensive margin (e.g., growth of already existing export lines), along the lines of the work by Bahar and Rapoport (2018). In either case, the results show that the elasticity of exports to return workers ranges from 0.08 to 0.14, depending on the transformation of the left hand side variable used (and thus whether zeros are included or not).

Columns 4, 5 and 6 present the analogous 2SLS estimates. For those columns we also report the Kleibergen-Paap F statistics which measures the strength of the first stage. The Kleibergen-Paap F statistics is the right measure to look at when standard errors are not assumed to be i.i.d., as in our case. The high magnitude of the F statistics in all specifications imply that we can reject the possibility of weak instrumentation.

While there is no technical problem with having a large F statistic, the large value in the 2SLS estimations might raise some concerns surrounding the exogenous variation that is being exploited and whether the fit is "artificial". Given that both treatment and instrument are originally defined for each German 3-digit WZ industry code, there might be concerns that our own concordance to SITC 4-digit codes might be the main source of the fit between the two (see Online Appendix Section 4.10.2 for details on the concordance between the two). However this is not the case. The first stage of our regression maintains its explanatory power even when using 3-digit WZ industry codes, our overall results remain qualitatively similar, in spite of using much fewer observations and reducing the variation on the dependent variable (see Online Appendix Section 4.10.7).

The elasticities estimated through 2SLS are positive, statistically significant and qualitatively similar to the OLS results but the point estimates are larger in magnitude. Yet, the standard errors are also larger, so that we cannot reject the hypothesis that the OLS and the 2SLS estimates are equal. Given the setting of the natural experiment, and the use of an instrumental variable, we interpret these results as causal. Thus, based on the 2SLS results, we find that Yugoslavian industries that received 10 percent more return migrants from Germany (that worked in those same industries), exhibited higher exports by 1.1 to 2.4 percent during the period of the study.

In the presence of self-selection of workers we would expect results to decrease in magnitude once we instrument. What may explain why the 2SLS estimates are in fact larger than the OLS ones? One possible explanation is that 2SLS uses variation in the treatment that is disproportionately coming from refugees allocated to areas in Western Germany (and specifically southern German states) where the most productive firms (and workers) are located . Results suggests that effects of resettlements are larger than for the average refugee, although we cannot reject the hypothesis that our OLS and 2SLS point-estimates are the same.

Overall, however, it is important to acknowledge that due to data limitation our treatment does not capture the exact number of refugees that indeed returned to their home countries (as explained in Section 4.6.1), but proxies for it. Unfortunately, we lack data to assess to what extent these workers

joined the former Yugoslavian countries' labor force upon their return and in whether they joined the same industries for which they worked while in Germany.¹⁴⁹ In this sense, akin to the concept of "intent to treat" estimators in the program evaluation literature, our results present a lower bound of the actual treatment effect.

Note, too, that our estimations use as the dependent variable exports from former Yugoslavian countries to the rest of the world, excluding Germany. In that sense, we argue that our results are not explained by possible reductions of fixed costs of exporting to Germany caused by migrant networks.¹⁵⁰

The table also reports estimates for β^C , related to our "convergence control". If convergence is something to be worried about (in terms of that process, and not the migrants driving our results) then this control will take care of that possibility. Note that the estimates for β^C are statistically insignificant, and therefore there is not much to say about the role of convergence based on them. In addition, to address this concern we also perform a number of placebo tests in Section 4.7.2 where we show that our treatment does not explain the same trends we see in other eastern European countries with a similar export structure than that of Yugoslavia in 1990.

A remaining plausible concern is whether our results are driven by other spurious correlations, such as pre-trends, scale effects or global demand driving our results. In the following sections we perform a number of tests to rule out other possible explanations. For instance, we look at pre-trends in the following subsection and find no evidence that industries with and without returning refugees behaved differently before the resettlement. Below, we also address concerns that our results are driven by scale effects – more workers returning resulting in larger export values – by showing that the effect we find is no different for industries based on their capital intensity level. In addition, we also show that our results are consistent when using industry-level productivity measures based on Costinot et al. (2012). Finally, we find that our results are driven by workers in occupations intensive in particular occupations, which tend to be just a small share of all workers (see Section 4.10.10). When it comes to global demand, our results –as we show below– are robust to controlling for changes in the relative size of each export sector at a global scale.

As an additional note, it is worth mentioning that when expanding this exercise to all countries –at the expense of weakening our identification strategy– we find our results to be externally valid (see Online Appendix Section 4.10.12).

4.7.1 Pre-trend and Event-study Estimation

Can our results be explained by a previous trend in exports? We explore this first by estimating the same specification but this time over the period 1985 to 1990, keeping the same treatment defined for

¹⁴⁹Though, as detailed in Section 4.4, historical accounts confirm that about 75 percent of Yugoslavian refugees in Germany returned to the former Yugoslavia.

¹⁵⁰Online Appendix Section 4.10.4 presents results for the main specification using as dependent variable exports to the rest of the world *including* Germany. As expected, the point estimates are (between 2 to 10 percent) larger.

years 1995 to 2000. OLS results are presented in Table 33, and in this case the estimates for β^{DID} are either non-significant or negative and very small.¹⁵¹ There is not enough evidence to conclude that the levels were different for treated and non-treated industries before the war.

Table 33: DID estimation, previous trend

Dependent variable: $exports_{p,t}$			
	(1)	(2)	(3)
	lnexp	lnexpplus1	asinhexp
treat2000 \times after1985	0.0282 (0.032)	-0.1002 (0.057)*	-0.1110 (0.060)*
N	1428	1572	1572
Adj R2	0.90	0.84	0.84

This table shows result of the estimation for specification (36) using different monotonic transformations for $exports_{p,t}$ in each column. The estimation uses years 1985 and 1990. All columns include product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Second, given the availability of exports data across several years, we turn to estimate the effect of return migration in a event study setting. To avoid noise in the estimation, we take 5-year averages for the dependent variable and estimate β^{DID} for 6 different periods, from 1985-1989 to 2010-2014. To do this, we simply re-estimate specification (36), this time substituting the dummy $after_t$ for several dummies, each one signaling a 5-year period, along the lines of Autor et al. (2003).¹⁵² In this multi-period setting, α_t are 5-year period fixed effects, and the product fixed effects η_p are maintained, allowing for product-specific intercepts. Table 34 reports the OLS and the 2SLS estimation using the instrumental variable described above (note we have in this setting four endogenous variables and four instrumental variables, which correspond to the treatment and the instrument multiplied by four different period dummies for periods post-1995). Every column includes the convergence control discussed in the previous section, though it is not being reported. Naturally, the number of observations in this sample is much larger than before, as it includes 6 data points for each of the 786 products totaling up to 4716 observations (except

¹⁵¹Note that we don't present 2SLS results because when using exports in a period previous to the treatment, there is no rationale for instrumenting the treatment. We also do not include the convergence control, for similar reasons.

¹⁵²We estimate the following equation:

$$exports_{p,t} = \sum_{t=1}^6 \beta_t^{DID} treat_p \times period_t + \beta_t^C \sum_{t=1}^6 DEUexpshare1990_p \times period_t + \eta_p + \alpha_t + \varepsilon_{p,t}$$

where $period_t$ is a binary variable that takes the value 1 for each corresponding 5-year period.

for the first column where observations where $exports_{p,t} = 0$ are excluded).

Table 34: Event study 5-year periods

Dependent variable: $exports_{p,t}$						
	OLS			2SLS		
	lnexp	lnexpplus1	asinhexp	lnexp	lnexpplus1	asinhexp
treat2000 × Period 1985-1989	-0.0422 (0.025)*	-0.0562 (0.040)	-0.0571 (0.042)	0.0250 (0.044)	0.0656 (0.065)	0.0666 (0.067)
treat2000 × Period 1995-1999	0.0299 (0.024)	0.0419 (0.039)	0.0420 (0.041)	0.1258 (0.047)***	0.2002 (0.074)***	0.2028 (0.077)***
treat2000 × Period 2000-2004	0.0520 (0.029)*	0.0990 (0.048)**	0.1006 (0.050)**	0.1467 (0.051)***	0.2724 (0.087)***	0.2771 (0.091)***
treat2000 × Period 2005-2009	0.0516 (0.030)*	0.1115 (0.054)**	0.1127 (0.056)**	0.1486 (0.057)***	0.2724 (0.094)***	0.2756 (0.097)***
treat2000 × Period 2010-2014	0.0710 (0.034)**	0.1675 (0.058)***	0.1702 (0.060)***	0.1209 (0.061)**	0.2836 (0.101)***	0.2883 (0.104)***
N	4585	4716	4716	4585	4716	4716
r2	0.85	0.80	0.79	0.85	0.80	0.79
KP F Stat				87.47	93.69	93.69

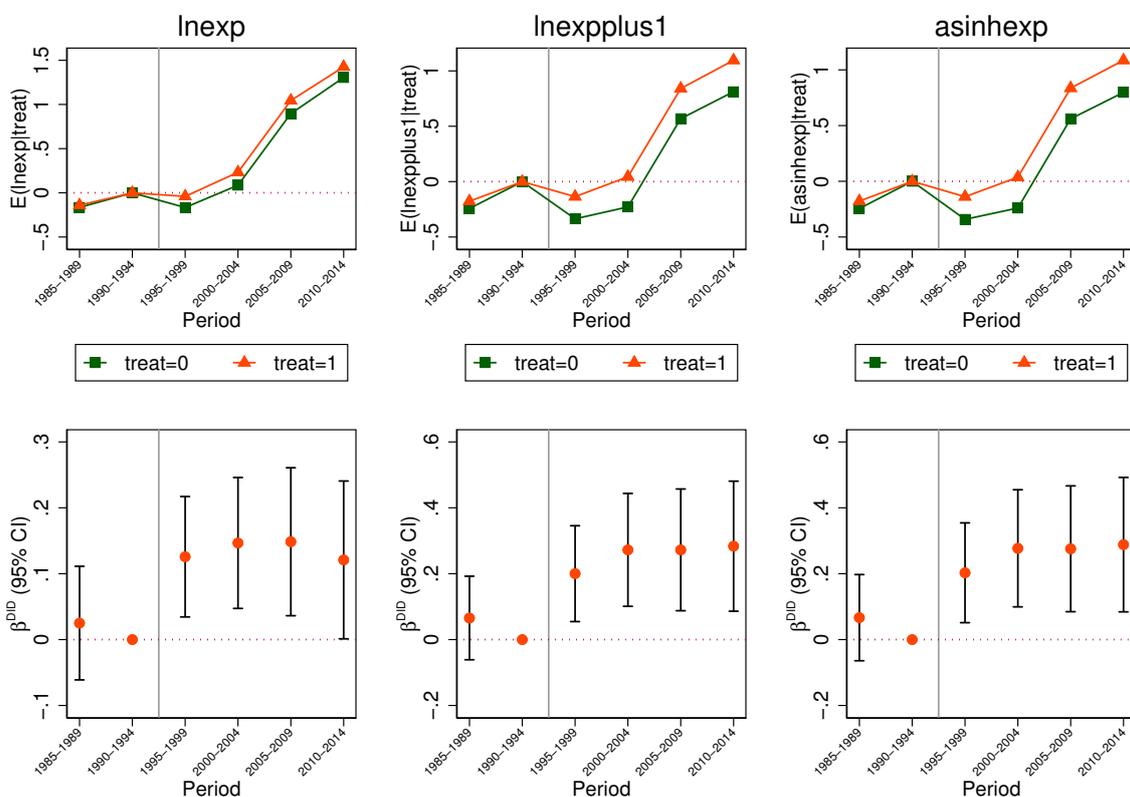
This table shows result of the estimation for specification (36) using different monotonic transformations for $exports_{p,t}$ in each column. It estimates the treatment across different 5-year periods. All columns include the convergence control, as well as product fixed effects and 5-year period fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Across all different dependent variables, our instruments are relevant as reported by the KP F statistic, though much smaller in magnitude than those reported in the main difference-in-differences specification. The estimations use the period 1990-1994 as the base for the estimation and is thus excluded, though all estimators should be interpreted as relative to that 5-year period. In that sense, across columns, we see that the treatment is zero or negative but barely statistically significant as compared to 1990-1994 in all of the periods before the actual treatment occurs (e.g., in periods 1985-1989 and 1995-1999). The OLS and 2SLS results indicate that the value of the elasticity are positive and statistically different from zero for every transformation of the dependent variable starting in the period 2000-2004, relative to 1990-1994 (and we see some of the effect starting already in 1995-1999 in the 2SLS results). When looking at OLS results, the elasticity in the first column is estimated to be 0.05 in the 2000-2004, after most migrants had returned. The same elasticity increases to 0.07 in period 2010-2014. In the other two columns, between 2000 and 2004, the elasticity is estimated to be around 0.1, almost double than that in Column 1. This strengthens our previous finding that the effect is stronger when we take into account the extensive margin. The elasticity grows up to 0.17 in the two last periods along the lines of our results from Table 32. The 2SLS point estimates in columns 4 to 6 are same in sign and slightly larger in magnitude to the OLS, consistently with what discussed above. They also show that the effect in the post-treatment period remains positive and significant relative to the period 1990-1994 beyond the period 1995-1999, when when most refugees returned (with the exception of column 5 where the effect seems not to be statistically significant in the period 2010-2014).

These findings suggests an important result: the marginal effect of return migration on the emergence of new exports becomes stronger over time. These results are summarized in Figure 31, which shows in the upper panel the evolution of the expected value of exports (across our three different measures) by 5-year periods for two groups of products using the 2SLS estimates: those for which the value of $treat_p$ equals 1, and the second group is those for which $treat_p = 0$. The figures in the lower panel show the difference between the two groups, and it can be seen how the effect becomes positive and statistically significant starting in the period where the refugees start returning, 1995 to 1999. Note that, based on the standard errors (as measured by the whiskers representing 95 percent confidence intervals), we cannot reject the hypothesis that the trends for both groups in periods before 1995 are statistically the same.

Figure 31: Event study (2SLS), 5 year periods



The top figure plots exports over time for two groups: products for which $treat_p = 1$, and products for which $treat_p = 0$. The dependent variable is the 5-year average of exports (each column uses a different linear transformation and the period 1990-1994 is used as the base year). The bottom figure plots estimates for β^{DID} for each 5-year period, which corresponds to the difference between the two groups of industries plotted in the top figure. The results are estimated using 2SLS and include the convergence control. 95% confidence intervals for the estimation are represented by the whiskers.

The figure also reveals that, the average industry experiences a drop in export performance during the war –as compared to the base period. Treated industries recover and reach back their 1990-1994 level in period 2000-2004 (or even earlier, when using the log transformation), while non-treated ones recover later on. In that sense, part of what our effect is capturing in the first few post-conflict years is

that treated industries, on average, shrunk less than non-treated ones.

4.7.2 Robustness and Heterogenous Effects

Foreign Direct Investment

As migrants could also reduce transaction costs and facilitate foreign direct investment (e.g., Javorcik et al., 2011; Kugler and Rapoport, 2007), we control for the stock of FDI from Germany (in million €) to the combined Yugoslavian countries in the specification.¹⁵³ We do so in order to rule out the possibility that the increase in exports is driven by the inflow of FDI in the same industries the migrants worked at while in Germany. Why would we want to rule out this possibility? Actually, we don't necessarily want to rule it out, as FDI inflows are one potential mechanism through which migrants can induce a productivity shift in their industries back home. However, by including this control we simply rule out FDI as mechanism, and instead focus on the idea that migrants, regardless of their ability to bring in investment, can explain changes in the composition of the export basket of their home countries.¹⁵⁴ Yet, even if part of the effect we document is driven by migrants being able to attract capital (or invest their own) resulting in the emergence and growth of export industries, this also reflects the existence of inputs complementary to capital (e.g., the workers themselves and their skills) in the economy. Table 35 presents the results of the estimation controlling for German FDI in the Yugoslavian countries.

Table 35: DID, controlling for German FDI

Dependent variable: $exports_{p,t}$	OLS			2SLS		
	lnexp	lnexpplus1	asinhexp	lnexp	lnexpplus1	asinhexp
	treat2000 × after2005	0.0856 (0.038)**	0.1396 (0.063)**	0.1415 (0.065)**	0.1365 (0.051)***	0.2750 (0.091)***
germanexpshare1990 × after2005	10.2621 (4.500)**	6.3892 (6.737)	6.2181 (6.937)	5.0949 (5.198)	-8.1044 (12.746)	-8.7197 (13.187)
lnfdi	-0.1619 (0.066)**	-0.2723 (0.127)**	-0.2770 (0.131)**	-0.1635 (0.066)**	-0.2770 (0.127)**	-0.2819 (0.132)**
N	1496	1572	1572	1496	1572	1572
r2	0.86	0.82	0.81	0.86	0.82	0.81
KP F Stat				597.74	651.82	651.82

This table shows result of the estimation for specification (36) using different monotonic transformations for $exports_{p,t}$ in each column. The estimation uses average exports for years 1988 to 1990 in the initial period and average exports for years 2005 to 2007 in the end period. It includes as a control the German FDI stock in Yugoslavia by industry. The first three columns report results from an OLS estimation, while the last three columns report results from a 2SLS estimation. All columns include product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Overall our main results are robust to the inclusion of FDI stocks. We find, however, that FDI is negatively correlated with exports, which is puzzling. However, since the data for FDI stocks was

¹⁵³Our results are unchanged if we include instead a broader FDI figure from all countries (in million €), not only from Germany. See Online Appendix 4.10.5 for these results.

¹⁵⁴There is also the possibility that FDI is a "bad control", given that the expansion of the labor force with skills relevant to a particular industry, which is what our treatment measures, can also attract FDI into those same industries, and hence we exclude it from our main specification.

originally at the 2-digit level (see Section 5.5), there is little variation left in it after the introduction of product fixed effects.¹⁵⁵

Falsification Tests

In this subsection we aim to show further evidence that our results cannot be explained by other economic processes occurring at the same time, such as convergence, which might be correlated with migration flows. We do so by putting in place two falsification or "placebo" tests.

First, we check whether return migration to (the former) Yugoslavia can explain export changes in similar countries in terms of their export baskets in the baseline period. To do so, we re-estimate our main specification but using as dependent variable exports from countries other than Yugoslavia to the rest of the World (excluding Germany) with the same treatment (e.g., return refugees from Germany to Yugoslavia). We focus on countries with similar export similarity to Yugoslavia; in particular, those that have an export similarity index with Yugoslavia above the 75th percentile of the distribution in 1990, using the index proposed by Finger and Kreinin (1979).¹⁵⁶

We further limit the group countries that, on top of having a similar export structure, also are in eastern Europe and have a similar income per capita (within a 25 percent range) than that of Yugoslavia in 1990. The countries that satisfy those conditions are Bulgaria, Poland and Romania. However, migrants from these countries have presence in the German labor force distributed in similar industries as Yugoslavian refugees in the early 1990s. Therefore it is important for us to add as an additional control to our regression the stock of those countries' migrant workers in Germany in 1995 per each SITC 4-digit industry. This is because Bulgarian, Polish and Romanian emigrants in Germany could also explain changes in exports in their home countries (as shown by Bahar and Rapoport, 2018). The results of this exercise are visualized in Figure 32, based on one regression per country using OLS estimators. The treatment is statistically insignificant for all three countries across every specification. For comparison purposes, the figure also shows the original (OLS) results for Yugoslavia, which is positive and statistically significant.¹⁵⁷

¹⁵⁵Online Appendix Section 43 estimates the correlation between Yugoslavian exports and FDI stocks using the same regression setting, and finds a positive coefficient when not including product fixed effects, which is the sign we would expect in such relationship.

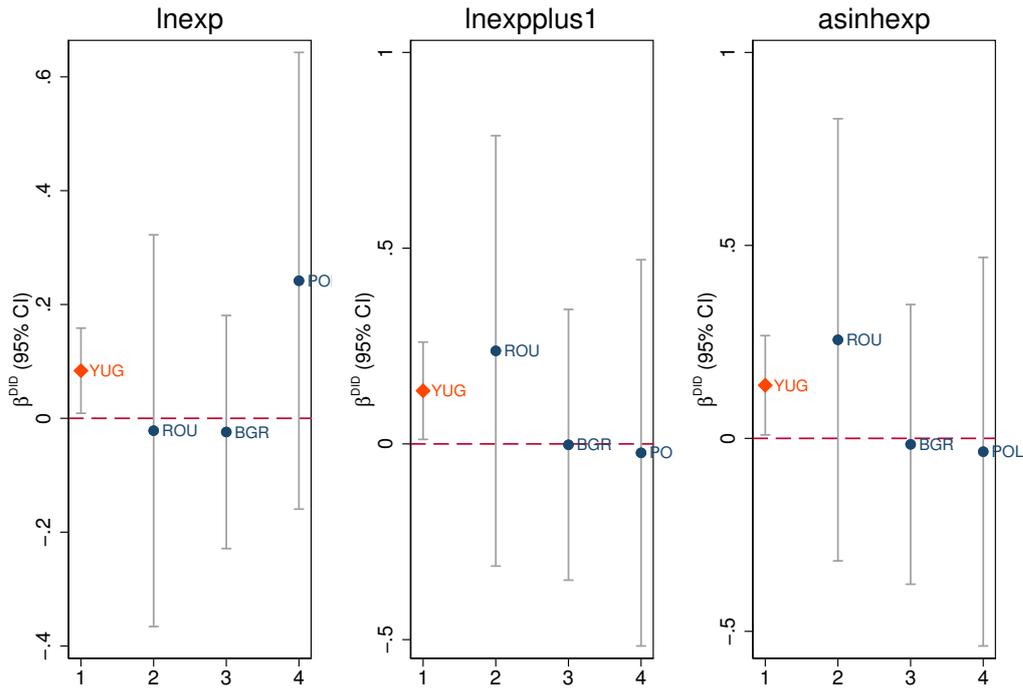
¹⁵⁶The export similarity index by Finger and Kreinin (1979) is constructed using the formula:

$$S_{c,c'}^{F\&K} = \sum_i \min(s_i^c, s_i^{c'})$$

where i represents industries, c and c' represent any two countries and s_i^c is the share of industries i in country c 's export basket. Hence, two countries c and c' that export the exact same industries in the exact same proportions would have an index equal to one.

¹⁵⁷Concerns can arise because the lack of significance could be driven by multicollinearity of the actual treatment with the additional control that we have added, namely the stock of migrant workers from those countries in Germany per industry in 1995. There are not many countries highly similar to Yugoslavia in 1990 in terms of their export basket, with little to no workers in Germany by 1995, except for only one: South Korea. Korea's export similarity with Yugoslavia in 1990 is above the 90th percentile of the distribution and Korea had less than 1000 migrant workers in Germany in 1995 based on our sample. Note, however, that already in 1990 Korea was very different from Yugoslavia in terms of other outcomes, such as income per capita, and thus not a very credible counterfactual. Nevertheless, when redoing this falsification test using Korea –without an additional control for the stock of Korean workers in Germany in 1995– we find an estimate that lacks statistical significance at conventional levels, which is a reassuring result.

Figure 32: “Placebo” test using exports from similar countries



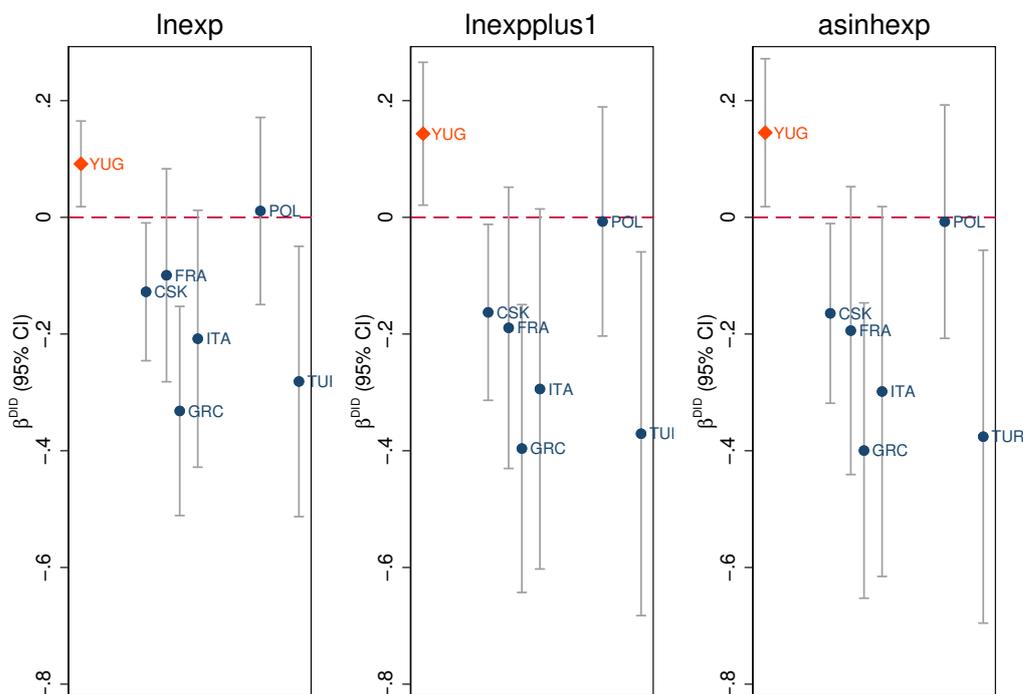
This figure plots coefficients of the estimation for specification 36 for each country, using different monotonic transformations. The estimation uses exports averaged over 1988 to 1990 for the initial year and exports averaged for 2005 to 2007 as the end year. The results are estimated using OLS, and include the convergence control, as well as the stock of migrant workers from each respective country in the case of Poland and Romania (and not in the case of Korea). 95% confidence intervals for the estimation are represented by the whiskers.

The second falsification test we run is somehow the mirror image of the first one and, essentially, it addresses an alternative explanation. The placebo test aims to rule out that the changes we see in Yugoslavia are driven by the shrinkage of the same industries in Germany (shedding workers, migrants among them), as part of a global general equilibrium of structural transformation process. Thus, the test asks whether the number of Polish workers, for example, that were part of the German labor force in 1995 but not in 2000, explain changes in the Yugoslavian export basket. To do this we focus on the five countries other than Yugoslavia with the largest changes in the stock of migrant workers in the German labor force between 1995 and 2000. These countries are, in order, Turkey, Greece, France, Poland and Italy. Specifically, we estimate the same main model using Yugoslavian exports as the dependent variable, with two differences. First, the treatment is the number of migrant workers from these five countries that were in the sample in 1995 and left by 2000 by SITC 4-digit industry; and second, under the same logic as the previous placebo test, we control for the stock of Yugoslavian migrant workers in Germany in 1995 for each SITC 4-digit industry.

The results are visualized (based on one regression per country) in Figure 33. The figure shows that the treatment is statistically insignificant for all countries across every specification. For comparison purposes, again, the figure also shows the original (OLS) results for Yugoslavia, which is positive and

statistically significant.

Figure 33: “Placebo” test using return migrants to other countries



This figure plots coefficients of the estimation for specification 36 for each country, using different monotonic transformations. The estimation uses exports averaged over 1988 to 1990 for the initial year and exports averaged for 2005 to 2007 as the end year. The results are estimated using 2SLS, and include the convergence control, as well as the stock of workers from each respective country. 95% confidence intervals for the estimation are represented by the whiskers.

Comparative Advantage

Finally, we reestimate our specification this time using the country-product specific measure of productivity or comparative advantage for Yugoslavia, $\Phi_{p,t}$, estimated following Costinot et al. (2012) and the application by Leromain and Orefice (2014).¹⁵⁸ Our alternative specification is then:

$$\Phi_{p,t} = \beta^{DID} treat_p \times after_t + \beta^{fdi} fdi_{p,t} + \eta_p + \alpha_t + \varepsilon_{p,t}$$

Following the previous results, we use the same two different monotonic transformations for $\Phi_{p,t}$ (given that since there are no zero-values we skip the $\log(\Phi_{p,t} + 1)$ transformation). Results are presented in Table 36. As can be seen, our results are robust to using this measure as the dependent variable in terms of sign and significance.

¹⁵⁸According to Costinot et al. (2012), $\Phi_{p,t} = e^{(\phi_{p,t}/6.53)}$, where the figure 6.53 is their estimation of the elasticity of (adjusted) bilateral exports with respect to observed productivity, and $\phi_{p,t}$ is estimated as the country-product specific productivity parameters for Yugoslavia using the following specification and using the complete matrix of bilateral trade (where Yugoslavia is one of the c countries in the dataset):

$$asinh(exp_{c,c',p,t}) = \phi_{c,p,t} + \Omega_{c',p,t} + \Psi_{c,c',p} + \varepsilon_{c,c',p,t}$$

In the specification $exp_{c,c',p,t}$ is the export value from country c to country c' of product p in year t , $\phi_{c,p,t}$ is a exporter-product-year fixed effect, $\Omega_{c',p,t}$ is a importer-product-year fixed effect, $\Psi_{c,c',p}$ is a exporter-importer-product fixed effect and $\varepsilon_{c,c',p,t}$ is the error term.

Table 36: DID, Costinot et al. (2012) measures in LHS

Dependent variable: $\Phi_{p,t}$, based on Costinot et al. (2012)				
	OLS		2SLS	
	lnxp	asinhxp	lnxp	asinhxp
treat2000 \times after2005	0.0205 (0.002)***	0.0162 (0.002)***	0.0246 (0.003)***	0.0195 (0.002)***
germanexpshare1990 \times after2005	4.2863 (1.731)**	3.7160 (1.466)**	3.8478 (1.551)**	3.3590 (1.317)**
N	1572	1572	1572	1572
r2	0.91	0.90	0.91	0.90
KP F Stat			723.55	723.55

This table shows result of the estimation for specification (36) using different monotonic transformations for $\Phi_{p,t}$ in each column. $\Phi_{p,t}$ is a measure of comparative advantage estimated following Costinot et al. (2012). The first two columns report results from an OLS estimation, while the last two columns report results from a 2SLS estimation. All columns include product fixed effects and year fixed effects. All columns include product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Effect for Industries with "Early" Returnees

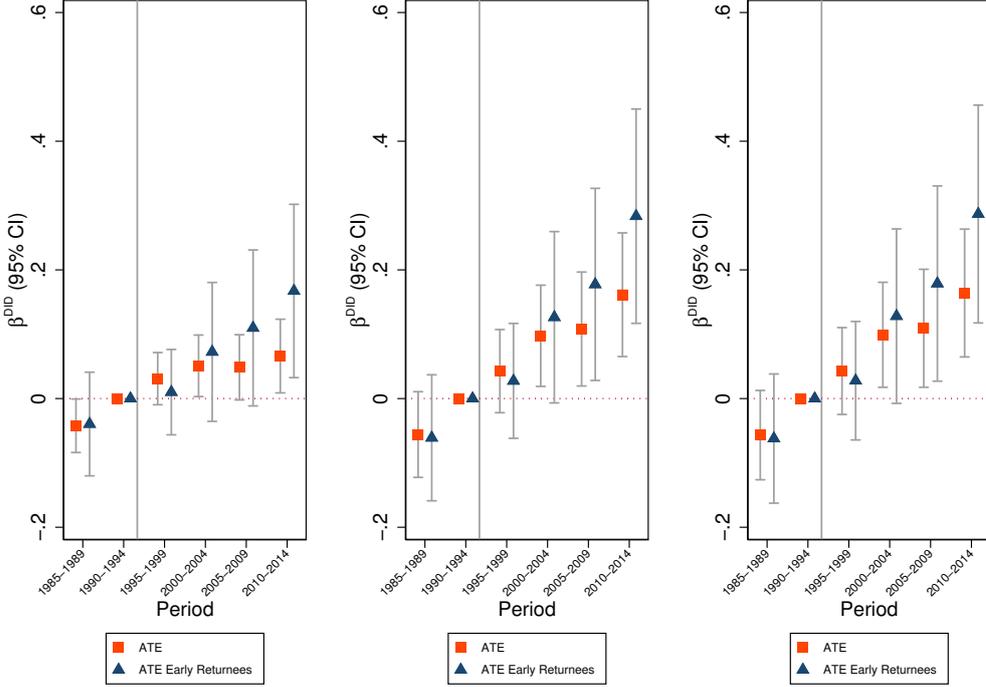
The nature of our dataset allow us to identify for each refugee worker the years of arrival to and exit from the German labor force. We know that most Yugoslavian refugees in Germany started returning to their home countries after 1995, as soon as their Duldung status expired. Thus, we are able to test whether there is a differential effect for industries for which most migrant and refugee workers returned earlier than others. Even if suggestive, this test could be very instructive for one particular aspect of our exercise: that the effect is indeed driven by returning workers, and less so by all refugees including those who did not return.

To test for this, we replicate the same event study documented in Table 34 and visualized in Figure 31, this time adding new terms to the specification to investigate whether industries with many workers having returned in the earlier part of the period (before 1996, inclusive) experienced a larger marginal effect early on, as compared to the rest of the industries. To do this, we include in the event study, on top of the regular $treat_p \times period_t$ variables for all 6 periods from 1985-1989 to 2010-2014 (again, where 1990-1994 serve as the baseline), a triple interaction $treat_p \times period_t \times earlytreat_p$, where $earlytreat_p$ is defined as 1 if industry p has a share of early returnees above the 90th percentile of the distribution, which corresponds to about 65 percent. In other words, in those industries, at least 65 percent of all

migrants returned in 1996 or before, while the rest returned between 1997 and 1999.¹⁵⁹

The results are visualized in Figure 34 (based on OLS estimations). The squares represent the average treatment effect, (ATE) for all industries in each period, whereas the triangles represents the ATE for the "early" treatment industries in each period.¹⁶⁰ Note that there is one square and one triangle for each one of the six 5-year periods.

Figure 34: Treatment effect for industries with high share of early returnees



This figure below plots the average treatment effect (ATE) in each 5-year period represented by squares, and the ATE for industries for which the treatment is composed by over about 65 percent (or 90th percentile of the distribution) of worker returnees who returned in 1996 or before. Both ATEs are presented for each 5-year period from 1985-1989 to 2010-2014 (with 1990-1994 serves as baseline). The results are estimated using OLS and include the convergence control. 90% confidence intervals for the estimation are represented by the whiskers.

The figure consistently shows that there is a "premium" for early treatment industries: the point estimates for the industries identified as having mostly early returnees tend to perform better earlier in time than the rest of the industries. Even though statistically we cannot reject the effects are different from each other, we can reject they are different from zero in the later periods. This suggestive evidence further supports the idea that it is returnees shaping the dynamics of export performance back in their

¹⁵⁹In other words, we estimate the following specification:

$$\begin{aligned} exports_{p,t} = & \sum_{t=1}^6 \beta_t^{DID} treat_p \times period_t + \sum_{t=1}^6 \beta_t^{DID,early} treat_p \times period_t \times earlytreat_p \\ & + \beta_t^C \sum_{t=1}^6 DEUexpshare1990_p \times period_t + \eta_p + \alpha_t + \varepsilon_{p,t} \end{aligned}$$

Where we define the average treatment effect (ATE) for industries with many early returnees (e.g., $earlytreat_p = 1$) for period t as $\beta_t^{DID} treat_p + \beta_t^{DID,early} treat_p$, while for the rest of the industries $\beta_t^{DID} treat_p$.

¹⁶⁰The ATE for the early treated industries is the sum of the estimated coefficient for the term $treat_p \times period_t$ plus the one estimated for the term $treat_p \times period_t \times earlytreat_i$.

home countries, and not only defined by the total number of refugees including those who return later on or do not return at all. This test is an additional proof that our results are not driven by convergence dynamics.

Beyond Exports: Focusing on the Bosnian Economy

So far, the results show a positive relationship between returning refugees and exports. In this section, we focus on Bosnia and Herzegovina, for two reasons. First and foremost, this is the only former Yugoslavian republic for which we were able to collect sectoral data other than exports prior to 1991, such as number of firms and number of employees. Second, it is the country with the largest number of refugees hosted by Germany during the war.

For this exercise, we digitized sectoral data from different official sources in order to construct our panel. The source of our pre-war data collection comes from the Statistical Almanac (*Statistički Godisnjak*) put together by the newly established Republic of Bosnia and Herzegovina. The government brought together data collected by multiple administrative units on a wide range of issues (e.g., health, demographics, municipal elections, production, exports and more) to provide a comprehensive picture of the social and economic state of the country. From this survey we collected sectoral data for export values, the number of firms, and number of workers for the year 1990.¹⁶¹ For the period after the war, we collected corresponding data for the year 2010 (the first post-war year for which such data is available) from the Structural Business Statistics reports (*Strukturne Poslovne Statistike*) published by the Statistical Agency of Bosnia and Herzegovina (*Agencija za Statistiku BiH*).

We match these data with the main treatment used in our baseline results. In addition, we construct and match two other variations of the treatment meant to capture more precisely the number of Bosnian returning refugees (as opposed to Yugoslavians in general). The first one of these corresponds to Yugoslavian refugees who joined the German labor force between 1991 and 1995 and left sometime between 1996 and 1999. Since returning to Bosnia was not safe before the end of 1995, most of the returning refugees during 1996 to 1999 are likely to be of Bosnian origin; in addition, most Bosnians who returned home did so after 1995. In other words, we exclude from that treatment those Yugoslavian workers who disappear from the German Social Security records prior to 1996. The second additional treatment is somewhat more restrictive as it counts workers who had been registered as having a Bosnian passport at some point during their stay in Germany.¹⁶² Consistently, it counts workers who had arrived sometime between 1991 and 1995 and had left by year 2000.

Bosnian data for both periods are defined using the 2-digit NACE classification, which is equivalent

¹⁶¹In fact, we collect sectoral data before the war for years 1988 to 1990, and consistently with our main estimation we use average exports per sector between 1988 and 1990 for the "before" period. Note that, differently from the main exercise, the data does not allow us to exclude exports to Germany.

¹⁶²While some of these workers entered the German labor force as Bosnian nationals, others obtained Bosnian citizenship during their stay in Germany, and reported it. For constructing this alternative treatment, we consider individuals with Bosnian nationality at any point during their stay.

to German WZ 93. This makes it possible for us to link data for both the left-hand side (i.e., sectoral outcomes) and the right-hand side (i.e., treatment) without need to use correspondence guidelines.¹⁶³ We limit our sample to those sectors for which we have information on the number of workers and of firms both in 1990 and 2010. This results in a sample of 20 sectors, and thus 40 observations in total. Consistently with previous estimations, we transform all the variables using the inverse hyperbolic sine and based our estimations on Specification (36).

Table 37 presents the OLS results using all three treatments and including the convergence control. We only show OLS results because we have a very small number of observations, and refrain from performing 2SLS estimations on such a small sample. Hence, these results are suggestive only.

¹⁶³We do still apply concordances to construct the convergence control, which is based on German exports in 1990 using the 4-digit SITC trade data used in the rest of the paper.

Table 37: DID, outcomes in Bosnia

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable: $asinh(outcome_{s,t})$	$asinhexp$	$asinhexp$	$asinhexp$	$asinhfirms$	$asinhfirms$	$asinhfirms$	$asinhworkers$	$asinhworkers$	$asinhworkers$
$treat2000 \times after2010$	0.3497 (0.143)**			0.5847 (0.144)***			0.2896 (0.107)**		
$treat9699 \times after2010$		0.3380 (0.148)**			0.5509 (0.146)***			0.2720 (0.105)**	
$treatBIH2000 \times after2010$			0.3632 (0.140)**			0.5977 (0.124)***			0.3057 (0.105)***
$germanexpshare1990 \times after2010$	14.8751 (9.524)	14.8639 (9.596)	14.5866 (9.392)	0.3211 (11.528)	0.4953 (11.659)	-0.0319 (11.306)	5.3615 (10.042)	5.4591 (10.095)	5.0572 (9.928)
N	40	40	40	40	40	40	40	40	40
Adj R2	0.92	0.92	0.92	0.85	0.85	0.86	0.66	0.66	0.67

This table shows result of the estimation for specification (36), focusing only on outcomes for the Bosnian economy, using variation of 2-digit sectors. Columns 1-3 uses exports as the dependent variable, columns 4-6 uses number of firms, and columns 7-9 uses number of workers, all transformed using the inverse hyperbolic sine. All columns include sector fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

** $p < 0.10$, *** $p < 0.05$, **** $p < 0.01$

Column titles correspond to the outcome variables used in each regression. Columns 1 to 3 present the estimates when using Bosnian exports as the dependent variable. The point estimates are highly significant and are around 0.34-0.36. While the point estimates are not directly comparable to our baseline results given the different aggregation levels, it is reassuring, and somewhat telling given the more precise and intense treatment, that the point estimates are more than double in magnitude compared to baseline. In Columns 4 to 6 and 7 to 9, respectively, we use the number of firms and the number of workers as outcomes variables. We see that returnees are associated with an increase in both the number of workers and number of firms in each sector. These estimates are also statistically significant.

We see importance in these results as we are able to show that the effect of return migration goes beyond exports and manifests itself for other meaningful outcomes as well. Moreover, point estimates using the treatment with Bosnian passports give larger elasticities and have smaller standard deviations for all three outcomes (exports, number of firms and number of workers) and are thus more precise (although the coefficients for the three treatments are not statistically significant).

Looking carefully at these elasticities, baring the fact that these are suggestive results, we also can see some interesting patterns. For example, the fact that the elasticity for exports is larger than the one for number workers, implies that treated industries end up with a higher level of exports per worker, which would be consistent with higher labor productivity. One can also note that the elasticity for number of firms is smaller than the one for number of workers, implying that the treated industries have in the "after" period more firms that are, on average, smaller. We don't have a prior on what to expect in terms of firm sizes, but this result would suggest that many of these firms are new, and therefore, small.¹⁶⁴ Consistently, there is rich anecdotal evidence on the creation of new firms by returnees, some of which we compile in Online Appendix Section 4.10.1.

Effect by Geographic Regions

Table 38 presents results for the estimation of β^{DID} according to specification 36, where in each row the left hand side variable includes exports to a particular geographical region. Interestingly enough, this table suggests that our main results using $\log(exports_{p,t})$ (e.g., products that already existed in the Yugoslavian export basket) hold across all regions except the Middle East and North Africa (**meastnafr**) when looking at the OLS results, though when looking at the 2SLS estimators some of the point estimates remain similar in magnitude but lose statistical significance. Thus, for regions such as Western and Eastern Europe (**weurope** and **eeurope**, respectively) –traditional trade partners for Yugoslavian countries– the results imply that there was an increase in the intensive margin of exports. However, when including zero exports on the left-hand side (columns 2, 3, 5 and 6) results are estimated as positive and significant for North America (**namerica**), **meastnafr** and **lac** regions. In fact, the number of

¹⁶⁴Constructing these ratios and using them as left-hand side variables result in estimates that are consistent with our observation in terms of sign, but they lack statistical significance, and therefore we do not report them.

new products exported to these regions is significant. When looking at Yugoslavian exports to `lac`, for example, export values are zero for 421 products (out of 786) in 1990. That set of products with zero exports in 1990 is reduced to 224 in 2005. The same figures for `meastnafr` are 199 (in 1990) and 118 (in 2005). Thus, this implies that part of the effect we are finding when limiting the sample to those regions is driven by new exports to these two destinations.

Table 38: DID, by region of destination

Exports to...	OLS			IV		
	$\log(exp)$	$\log(exp + 1)$	$asinh(exp)$	$\log(exp)$	$\log(exp + 1)$	$asinh(exp)$
<code>weurope</code>	0.097**	0.102	0.099	0.153***	0.136	0.131
<code>eeurope</code>	0.293***	-0.233	-0.272	0.194**	-0.386	-0.434
<code>namerica</code>	0.176***	0.208**	0.206*	0.168*	0.399***	0.406***
<code>easiapac</code>	0.209***	0.111	0.101	0.004	0.055	0.047
<code>meastnafr</code>	-0.003	0.192**	0.204**	0.043	0.308**	0.324**
<code>lac</code>	0.197***	0.239**	0.237**	0.040	0.319**	0.324**
<code>ssa</code>	0.169***	0.202**	0.200**	0.106	0.188	0.187

This table shows result of the estimation for β^{DID} from specification (36), where each rows limits the importing countries to the different geographic regions. The first three columns report results from an OLS estimation, while the last three columns report results from a 2SLS estimation. All columns include product fixed effects and year fixed effects. Each group of results uses different monotonic transformations for $exports_{p,t}$ in each column. All columns include FDI as control, as well as product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We believe these results are particularly important, because they are consistent with the idea that following the return of these migrants back home, Yugoslavian countries started exporting new products to new destinations, implying that the effect we are finding reflects a structural change in the export activity of the country. Note that these results by sub-regions also mitigate concerns that our results could be driven by the effect of migration networks on transaction costs, given the small community of Yugoslavian migrants in Latin America and in the Middle East and North Africa.

Product Heterogeneity Analysis

Our main interpretation for the above results is that Yugoslavian returning refugees were able to increase the productivity of industries they worked in while in Germany thanks to the knowhow and experience they acquired abroad. In this section we proceed to rule out alternative explanations and do so by re-estimating specification (36), this time interacting the term $treat_p \times after_t$ with different product characteristics. First, we look at differentiated versus homogenous and reference-priced goods, using the definition of Rauch (1999) (i.e., a dummy variable). Second, we use the physical capital intensity level of each product, as defined by Nunn (2007) using data from the NBER-CES Manufacturing Industry Database of Becker et al. (2013). Third, we also include the upstreamness measure from Antràs et al.

(2012) which quantifies average distance from final use for each product.¹⁶⁵ Fourth, we use human capital intensity taken from Shirotori et al. (2010) to study whether there is a differential relationship between migrants and products with different knowledge intensities. These last three variables are continuous, and we standardize them to have zero mean and a standard deviation of one. We report only 2SLS results to avoid redundancy. Given that these variables have certain level of correlation in between them, our estimation includes all of them simultaneously (though results are robust to including them separately, too).

First, we address the possibility that the results are driven by the ability of migrants to lower trade transaction costs, making exports are more likely upon their return. This is what we check in the second row of Table 39, which interacts the treatment with a dummy indicating whether the product is differentiated. At first, we should not worry much about this possibility, given that our dependent variable already excludes exports to Germany. However, a concern remains if some of these migrants instead of returning to Yugoslavia migrated to third countries, and the increases in exports we are catching are induced by the decrease in transaction costs between Yugoslavian countries and, say, Austria or Belgium. However, as can be seen, the effect for differentiated products (those that are more likely to react to changes in trade transaction costs) is barely statistically different than zero, but most importantly for our purposes, we cannot reject the hypothesis that it is statistically different than the estimator in the first row, which corresponds to homogenous goods. Thus, there is no evidence to support an interpretation based on decreasing trade transaction costs.

Second, according to some of the most traditional trade models based on capital and labor endowments, our results could be driven by the fact that an inflow of workers into the economy could result into the export basket shifting towards labor intensive goods (Rybczynski, 1955). Yet, the results are not different for goods at different levels in the scale of capital intensity (denoted by `ki`), as seen in the third row of Table 39. In that sense, our results are not explained by scaled effects driven by a larger influx of workers into the economy.

Third, it is interesting to understand whether the treated industries are those with higher added-value as measured by the upstreamness measure (denoted by `upstr`), which proxies –to some extent– more complex production processes. We find, when looking at the fourth row, that indeed the effect is stronger for industries that are closer to final use (i.e., a lower levels of `upstr`).¹⁶⁶ Relatedly, if part of the results we are capturing have to do with the diffusion of knowledge, then we should expect some differential effects in terms of the knowledge intensity of the good. Interestingly, we find that return migrants explain more exports in products that are higher in the scale of knowledge intensity (as measured by human

¹⁶⁵The measures were provided to us in NAICS codes by the authors, we matched to SITC 4-digit codes following the approach described by Cuñat and Melitz (2012) in their footnote 24 and their subsequent documentation.

¹⁶⁶In Online Appendix Section 4.10.5 we test for an alternative explanation that this particular result could be consistent with. This is, migrant workers facilitating the imports of intermediate goods for the industries they worked in. We include a control term measuring total imports of intermediate goods in our specification, and our results are robust to this inclusion.

capital intensity, denoted by HCI), as seen in row five. It is important to notice that this last set of results could be interpreted as a Rybczynski effect in a model of trade that incorporates factor endowments such as human capital, knowledge-workers or certain skills other. Yet, when we refer earlier to Rybczynski we are thinking of the traditional model with labor and capital.

Table 39: DID, heterogeneity analysis

Dependent variable: $exports_{p,t}$			
	lnexp	lnexpplus1	asinhexp
TreatmentXAfter	0.0494 (0.065)	0.0767 (0.095)	0.0767 (0.098)
TreatmentXAfterXdiff	0.0858 (0.062)	0.1275 (0.069)*	0.1291 (0.070)*
TreatmentXAfterXki	0.0065 (0.025)	0.0152 (0.029)	0.0154 (0.029)
TreatmentXAfterXupstr	-0.0740 (0.028)***	-0.0636 (0.030)**	-0.0633 (0.030)**
TreatmentXAfterXhci	0.0708 (0.029)**	0.0943 (0.034)***	0.0953 (0.034)***
N	1180	1202	1202
r2	0.86	0.80	0.80
KP F Stat	112.72	118.13	118.13

This table shows result of the 2SLS estimation for specification (36), interacting the term $treat_p \times after_t$ with three variables indicating product characteristics: differentiated vs. homogeneous goods (top panel), capital intensity (middle panel) and human capital intensity (bottom panel). Each group of results uses different monotonic transformations for $exports_{p,t}$ in each column. All columns include product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.8 Mechanisms: Heterogeneous Effects by Workers' Characteristics

The idea that migrants can play a role in shaping the comparative advantage of countries is part of a growing literature that links migrants and their descendants to the diffusion of knowledge (e.g., Kerr, 2008; Choudhury, 2016; Hausmann and Neffke, 2016; Bahar and Rapoport, 2018; Kerr, 2018), and our results so far suggest this mechanisms could be a plausible one. If this were the case, we should be able to see stronger results when looking at migrant workers more suited to acquire and transfer knowledge.

This is what we explore in this section, where we study the role of different types of occupations in explaining changes in comparative advantage. Indeed, an important question that has been looked at recently in the labor economics literature is whether certain occupations, especially those intensive in managerial skills, are essential in fostering productivity (e.g., Bloom and Van Reenen, 2007; Bloom et al., 2012, 2013, 2019). We contribute to this literature by looking at whether our results are better explained by workers' skill levels and occupational patterns.

In particular, we expand Specification (36) and rewrite it as:

$$exports_{p,t} = \sum_i \beta_i^{DID} treat_{p,i} \times after_t + \beta^{fdi} fdi_{p,t} + \eta_p + \alpha_t + \varepsilon_{p,t}$$

Where each term $treat_{p,i}$ corresponds the total number of returning Yugoslavians in each category i in terms of workers' characteristics. All other terms remain the same as in Specification (36).

We present results using characteristics grouped in six different categories. First, instead of counting the number of returnees, we weight them with their last seen salary while in Germany (`wageKsm1`) in thousand € (euros). Higher salaries should not only reflect higher productivity, but the highest salaries within each industry is often believed to be a proxy for managerial tasks.¹⁶⁷

Second, skilled vs. unskilled workers based on their education levels. As unskilled we define workers without post-secondary education (`edulow`), and skilled as workers with education beyond high school, including vocational training, college degree or more (`eduhigh`). Since education does not devalue, we simply use the highest educational information attached to each worker at any point during the period of observation. To improve consistency of our variable, we correct missing values by using past and future values as developed by Fitzenberger et al. (2006).

Third, we distinguish migrants with occupations intensive in manual tasks (`taskmanual`) vs. occupations intensive in analytical and cognitive tasks (`taskanalytical`), using the classification provided by Dengler et al. (2014a), which formalizes German occupations into five task categories, similarly to Autor et al. (2003).¹⁶⁸

Fourth, we classify occupations as low skilled (`bf2lowskill`) and high skilled (`bf2highskill`) based on Blossfeld (1987) classification of professions. For example, high skilled occupations include managerial ones as well as professionals (i.e., engineers, lawyers, technicians, accountants, lab technicians), and low skilled occupations include drivers, carpenters, textile processing operatives, etc.

Fifth, we distinguish workers by the supervisory intensity of their occupation based on the German Qualifications and Career Survey (BIBB/BAuA) of 1999. In particular we use the workers' responses regarding their supervisory status¹⁶⁹ and assign to each occupation both the share of workers that self-

¹⁶⁷Using arbitrary wage cutoffs to identify managers in each industry results in very noisy measures. Thus, we use a continuous one.

¹⁶⁸Spitz-Oener (2006) first applied the task-based approach on German occupations based on survey data. The classification we use is based on year 2011.

¹⁶⁹Based on the answer to the question: "Do you have coworkers for whom you are the direct supervisor?".

report acting as supervisors (`svct1`) and the share of those that report the opposite (`svct0`).¹⁷⁰

Sixth, we distinguish workers based on whether they worked in the top 25 percent paying firms in terms of average wages (`fwaget251`), or in the bottom 75 percent paying firms (`fwaget251`). Typically, top paying firms are the most productive ones, by being able to attract the best workers and by innovating or adopting innovations that help workers be more productive.

Finally, we distinguish workers based on the average growth in their wage during their stay in Germany, as proxy for productivity improvements. We separate workers within each industry in two groups: workers with wage growth (based on the compound average growth rate, CAGR) below median (`wgrcagramd0`) and those with wage growth above the median (`wgrcagramd1`), based on all returnees in our treatment.

The summary of our results are presented in Table 40.¹⁷¹ Each column present results using a different monotonic transformation of the dependent variable, consistently with all previous results in the paper. Columns show the estimated value of β_i^{DID} for each of the constructed treatments belonging to each of the categories described above. We only present results using OLS, as we don't have instruments for more than one endogenous variable at a time.

¹⁷⁰Online Appendix Section 4.10.10 summarizes the values of these characteristics for each one of the occupations in our dataset, along with the number of workers in our sample in each occupation.

¹⁷¹See Online Appendix Section 4.10.11 for tables with all the estimations by group, including both univariate and multivariate regression. While there is multicollinearity, the relative size of the point-estimates remain consistent in univariate and multivariate regressions.

Table 40: DID, workers' education and occupations

β_{DID}	$\log(exp)$	$\log(exp + 1)$	$asinh(exp)$
migt1	0.084**	0.136**	0.138**
wageKsm1	0.148**	0.235**	0.238**
edulow	-0.279	-0.196	-0.192
eduhig	0.413***	0.381***	0.379***
taskmanual	0.022	0.071	0.072
taskanalytical	0.130*	0.140*	0.140*
bf2lowskill	0.007	0.059	0.061
bf2higskill	0.120*	0.124	0.125
svct0	-0.055	0.020	0.022
svct1	0.217	0.185	0.185
fwaget250	-0.033	0.022	0.025
fwaget251	0.155**	0.153*	0.152*
wgrcagramd0	-0.041	-0.036	-0.035
wgrcagramd1	0.131	0.184	0.186

This table shows result of the OLS estimation for specification (36) using treatments constructed by aggregating workers by groups based on their skills and/or occupations. The table presents OLS estimations. Each group of results uses different monotonic transformations for $exports_{p,t}$ in different columns column. All columns include the convergence control, as well as product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The first row replicates the main results using the total number of returnees per industry, for comparison purposes (migt1). Overall, based on the point estimates and statistical significance, our results show that our findings are stronger for workers with higher wage levels (row 2). They are also particularly driven by workers with higher educational attainment (rows 3 and 4), workers in occupations that are intensive in analytical tasks (as opposed to manual ones) and workers in skilled occupations (as opposed to unskilled ones). The results are also strongly driven by workers with occupations intensive in cognitive and analytical tasks, as opposed to manual ones (rows 5 and 6), as well as by workers in skilled occupations (rows 7 and 8). The results for workers in occupations intensive in supervision are

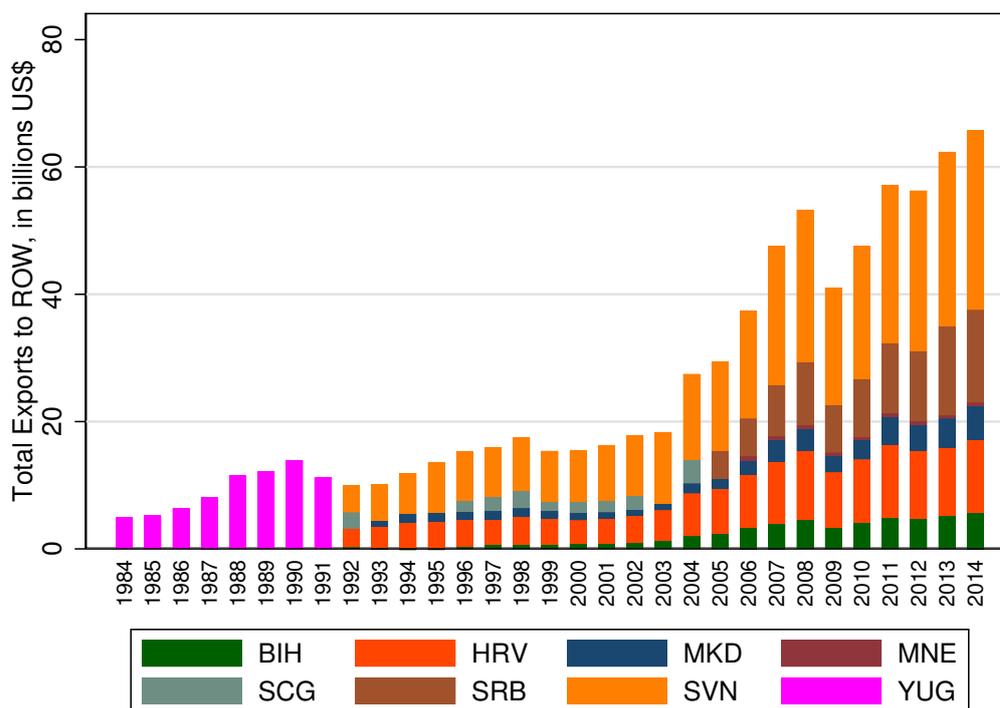
consistent, though not statistically significant, probably due to lack of variance (rows 9 and 10).

We also find that the results are particularly driven by workers who worked in the top paying firms during their stay in Germany (rows 11 and 12). We also find that workers for whom wages grew faster during their stay in Germany correlate with a higher export performance, though with no statistical significance (rows 13 and 14).

Note that the point-estimates are to be interpreted in terms of percentages, and thus, ultimately, the marginal effect of one worker belonging to each of the categories driving the results is much larger in relative terms than what the point estimate suggests (or 1000 euros of salary, in the case of row 2). This is because the types of workers driving the results are a smaller share when looking at the within-industry composition of workers in the sample. Figure 35 estimates the marginal effect of *one* migrant worker on exports using $\text{asinh}(\text{exports}_{p,t})$ as the dependent variable (results using $\log(\text{exports}_{p,t})$ and $\log(\text{exports}_{p,t} + 1)$ are qualitatively similar).¹⁷² The figure shows that, as compared to counting the number of workers (first bar), the marginal effect of each 1000 euros in salary for a worker is large and statistically significant. The figure also shows clearly that the marginal effect for workers with higher educational attainment is infinitely larger than for those with low educational attainment (given that the point estimate for the latter category is below zero). Similarly, workers in occupations intensive in analytical tasks are about 23 times more "effective" than those in occupations intensive in manual tasks. Workers in occupations that are considered skilled are about 35 times more effective than those in occupations considered unskilled. Workers in occupations for which supervision is more common are 25 times more effective than those in occupations where supervision is less common (though the difference is not statistically significant). Workers who were employed by the top 25 percent paying firms are, similarly to educational attainment category, infinitely more effective than those who worked in the bottom 75 percent. Similarly, a worker that experienced wage growth above the median value is much more effective than those with slower wage growth (though, again, not statistically significant). All in all, these results suggest that the size of the effect we document depends on who the workers are in terms of their skills, the characteristics of their occupations, as well as who did they work for and how successful in their jobs they were while abroad.

¹⁷²We compute this through a back-of-the-envelope calculation. We first compute what share of the the treatment represents one migrant worker (or 1000 euros for the case of wages), and multiply this share by the point estimate of β_i^{DID} .

Figure 35: Marginal effect by type of migrant



This figure plots the estimated marginal effect of 1 migrant returnee on exports from the home country based on the levels of migrants of each type in the sample. It uses $\text{asinh}(\text{exports}_{p,t})$ as the dependent variables. Whiskers represent 90 percent confidence intervals.

Our results for workers in occupations with managerial skills are consistent with the findings of Bloom et al. (2018), who find that management has an effect on firms' exports. By way of comparison, barring the differences in the way management is measured across our study and Bloom et al. (2018), the order of magnitude of both estimators are comparable. In the case of Bloom et al. (2018), they find that a one-standard-deviation rise in the management z-score is associated with 23% to 37% larger export revenues (for firms in China and the US, respectively). Our study shows that at the industry level, the elasticity of exports to an inflow of workers with foreign experience in occupations intensive in managerial skills lies between 0.14 and 0.5, depending on how we proxy for managerial skills.¹⁷³

The idea that a small number of workers can have such an important effect on exports of a whole industry in such little time might seem implausible at first, but some anecdotal evidence documented by others seems to strongly support that idea. For instance, Rhee and Belot (1990) and Easterly (2001), document the story of the success of the garment sector in Bangladesh. Between 1980 and 1986, the share of garments in Bangladesh's total exports rose from 0.5 to 28.3 percent. The unprecedented take-off of the garment export sector is often attributed to 130 Bangladeshi workers –only four of them in management positions– who spent eight months in 1979 working and being trained in Korea as part of

¹⁷³Relatedly, Bloom et al. (2019) find that the opening of a "million dollar plant" (MDP) in a US county results in an increase in productivity and in employment for incumbent local plants in the same county. This effect, they find, is driven by industries more likely to benefit from a managerial flows from the entry of the new plant.

an agreement between their company, Desh of Bangladesh, and the Korean firm Daewoo. The knowhow acquired by these workers seems to have been crucial in making Desh a highly successful exporter firm. Yet, perhaps more importantly, such knowhow eventually spilled over as workers moved to other firms or created new ones, contributing to the massive success of garment exports as one of Bangladesh's most significant export sectors.

In this context, we believe our findings pointing to productive knowledge and managerial knowhow as the main mechanisms driving the export dynamics, as well as the magnitudes of the coefficients we report, are aligned with other studies in the literature.

4.9 Concluding Remarks

The Balkan wars of the early 1990s created massive forced displacement from and within the former Yugoslavia. Most internationally displaced refugees ended up in Germany, where they could work under temporary protection status. A majority of them eventually returned home after the Dayton peace agreement of December 1995 and the repatriations that followed. We exploit this natural experiment, and the exogenous exposure to German industrial know-how and technology it created, to investigate the role of returning refugees in explaining the export performance of their home countries. Using confidential German social security data, we find that Yugoslavian exports performed significantly better during the post-war period in industries that returnees had worked in while in Germany.

Given that productivity is an underlying determinant of exports, we interpret these results as supporting the idea of migrants being drivers of knowhow and technology transfers between countries. This interpretation is backed by the fact that our results are particularly driven by industries that are knowledge intensive, and stronger when returnees are skilled, are in occupations intensive in analytical and cognitive tasks as well as in other managerial characteristics. In that context, our results are consistent with the literature linking productivity shifts and exports to improved managerial practices (e.g., Bloom et al. 2013; 2018; 2019).

Our results contribute to a burgeoning literature that emphasizes that migrants can serve as drivers of technology and knowledge diffusion resulting in productivity shifts, possibly reflected in several economic outcomes such as exports. To the best of our knowledge, we are first to find such evidence using a natural experiment – especially in a context of forced migration and return. In that sense, our results document how returning refugees, after having been integrated in their host economies' labor markets, can play a significant role in the post-conflict reconstruction of their home countries upon their return.

In terms of policy implications, our results are not meant to favor displacement, by no means. Rather, if displacement has occurred, it is in the best interest of the receiving country and the refugees to allow for full labor integration of the latter; not only for the obvious reasons, but also because of how their integration in labor markets can be a crucial determinant of the reconstruction of their home countries

upon their eventual return.

More generally, the ability of a worker to become more productive has to do with his or her accumulated experience and his or her ability to learn from others while on-the-job. Migration, therefore, is an important vehicle in the process of knowledge and technology transfer across locations. Better understanding this process and identifying channels through which these dynamics occur are important missing pieces in the literature, and an active part of our future research agenda.

4.10 Appendix

4.10.1 Anecdotal Evidence: Four Individual Stories

In 1999, only four years after having returned from Wolfsburg in Germany, Volkswagen's home town, Nijaz Hastor founded the Prevent group, currently one of Bosnia's largest companies. Prevent began manufacturing seat covers in the city of Visoko in Bosnia with a staff of 50, and has since diversified into yacht interiors, protective clothing, brake disks and fashion textiles. By 2016, the Prevent Group employed over 6,500 people and operated from about 15 different sites in Bosnia, exporting its products to a large number of different destinations across Europe and beyond. Hastor started his career working for a local firm supplying car parts in Sarajevo, but it is likely that the knowhow he acquired while working as an immigrant for the auto industry in Germany helped him build a world-class company able to manufacture high-end auto parts with high efficiency.

Almir Gvožđar is another example of a refugee who used his newly acquired skills and contacts to create his own company in a new industry. In 1996, following his return from Germany, Gvožđar invested all of his savings in a second-hand CNC machine tool and founded GAT Ltd in his family garage. Working alone, Gvožđar started producing motorcycle parts and selling them to his former employer in Germany, ABM Fahrzeugtechnik GmbH, a leading manufacturer of high-performance parts for the motorcycle industry, where he had worked as a technician during the war. Over the years, as the number of clients increased, business expanded as well. Currently GAT employs 65 people, operates from a facility of 1100 m² and exports motorcycle parts and medical instruments across Western Europe.

Refugees who were hosted in other countries also had similar experiences. For instance, Enes Kahrimanovic left BiH in 1991 and moved to Austria where he started working for the Plaspack Company, one of the largest manufacturers and distributors of nets, transparencies and advertising transparencies in the EU, a sector in which he had no previous experience. As Austria allowed its refugees to become permanent residents, Kahrimanovic continued working at the Plaspack following the peace treaty. While working at the Plaspack, Kahrimanovic realized that some of the intermediary products imported from rest of Europe can be produced in Bosnia. Over the years, Kahrimanovic worked on finding partners in Bosnia that could supply the imported pieces. As deals with local companies were struck, the Plaspack Company supplied more and more of its parts from Bosnia. Finally, Austrian owners of the Plaspack decided to start a production in Bosnia, and gave Kahrimanovic the full responsibility of both establishing and managing the company. In 2007, he founded Austronet in the city of Kozarac and with a staff of 5 and started manufacturing safety netting for tennis courts and protective netting for the construction industry. Today this company employs 72 people and exports 97% of its production.

Although individual initiative has been an important element for the transfer of knowhow, it was not indispensable as it can be seen in the case of the Kavat Shoe Factory workers. In the beginning of the

1990s shoemaking was a trade that was declining in Sweden and Kavat, a shoe manufacturer specialized in high-quality leather shoes was having difficulties finding skilled craftsmen. When Bosnian refugees from Travnik, a region specialized in the textile industry, arrived to Kumla where the Kavat factory is located, it didn't take long before they were recruited. By working for Kavat, Bosnians acquired skills in shoemaking and learned about modern equipment and technologies. Kavat was so satisfied with its Bosnian employees that when they returned to Travnik after the war it helped them establish a production facility and integrated it to its supply chain. Over the years, as the demand for Kavat shoes increased, the company felt the need to expand its production. The decision for location, as put by the company, came naturally. In 2009 Kavat opened a factory in Travnik and recruited its former employees. Today Kavat is an international company which manufacture about 400.000 pairs of shoes every year, out of which 350.000 are made in their factory located in Bosnia.

4.10.2 Details on Employment Data, Sample Construction and Variable Description

The data on migrant workers in Germany are based on records from the German social security system for the years 1975-2014 (IAB employment history (BeH), 2015) and comprise all persons employed subject to social security contributions, with the exception of self-employed and civil servants. The data contain information on nationality, education, occupation, industry, among others. For data privacy reasons our sample is restricted to 40% random draws of foreign nationals observed on June 30 of each year from 1975 to 2014 augmented by the employment history of each individual for our sampling period. The data was provided by the IT Services and Information Management (ITM) of the IAB. Missing information on educational attainment was corrected by ITM using information on past and future values (see Fitzenberger et al., 2006, imputation procedure IP1).

We keep all spells subject to social security contributions without specific tokens. Specific tokens are given to e.g. apprentices, employees in partial retirement, marginal part-time workers, seamen, or artists liable to social security. We keep one spell for per person-firm combination and focus on spells in tradable industries only. We use the nationality of the worker recorded at his or her first appearance in our database. The BeH contains information on the industry affiliation, but different classifications have been applied over time. Therefore, we use time-consistent industry codes developed for these data by Eberle et al. (2011). In particular we use the German classification WZ 93 which corresponds to the European classification of NACE Rev. 1. When matching German WZ 93 3-digit industry codes to 4-digit SITC product codes we apply correspondence tables provided by the United Nations Statistical Office and Dauth et al. (2014), which provide an unweighted concordance table. If the source 3-digit category applies to more than one 4-digit SITC target category we distribute workers according to the shares of German exports in 1995 based on 4-digit SITC categories for each year separately, along the lines of what implemented by Cuñat and Melitz (2012). Using this weighting scheme we use the implicit assumption that German export shares are a good proxy for employment of Yugoslavians across German at the product level (SITC 4 digits). This is something we cannot directly test for because our data on employment is at the industry level (WZ 93 3-digit). However, when looking at the correlation between German exports and employment share of Yugoslavian workers by industry (WZ 93 3-digit) we find it is reasonably high: for the years 1991 to 1995 is 0.72 and statistically significant. To comply with data privacy rules, the sample utilized herein is an anonymized version, sensitive values (between 1 and 19) of industry-period observations have been replaced with different moments of the distribution of the number of migrant workers. The number of cells affected varies by the level of disaggregation of worker characteristics.

The treatment variable for our main specification is constructed as follows. We keep all records of workers observed in the data for the very first time between 1991 and 1995 and for whom no record exists after 1999. For this we look at the total of records in our original data and can therefore rule out, that

workers in our treatment variable have had any other form of employment prior or after the respective cut-off dates. This comprises jobs in other industries (such as services), marginal employment, or any other form of employment which creates a notification to the social security services. Since a worker may have had two or more jobs in different industries, we assign each worker to his or her main job, before aggregating to the industry level. As main job we define the worker-firm-industry-occupation combination with the longest duration.¹⁷⁴

In some specifications in our analysis we disaggregate our data further according to several different worker characteristics. Our definition of the main job ensures for each dimension the sum each subcategory by industry (including potential missing values) adds up to the value in our main treatment variable. When distinguishing between skill groups, we define as unskilled, workers without post-secondary education and skilled as workers with education beyond high school (i.e., vocational training, college degree or more). Since education may change over time but cannot depreciate, we use the highest educational attainment. We also group workers according to the task content of their occupation. Occupations in our data are classified according to the German Classification of Occupations 1988 at the 3-digit level which comprises 334 different occupations. We distinguish between manual and analytically intensive tasks. Manual tasks are defined as manual (non-) routine tasks and as analytic tasks we classify analytical or interactive non-routine tasks. The classification is based on BERUFENET, which is, similar to O*NET, an expert's assessment of the tasks usually performed in a specific occupation. It covers originally about 3,900 different occupations and has been mapped to our classification codes by Dengler et al. (2014b). We use the classification for the year 2011. When classifying occupations by skill we categorize groups according to Blossfeld (1987). Low skilled occupation comprise agricultural, unskilled manual, unskilled services, and unskilled commercial and administration occupations. All other we define as high skilled.¹⁷⁵ We also distinguish workers by the supervisory intensity of their occupation based on the German Qualifications and Career Survey (BIBB/BAuA) of 1999. In particular we use the workers' responses regarding their supervisory status and assign to each occupation both the share of workers that self-report acting as supervisors and the share of those that report the opposite.

Furthermore, our employment data contain information on daily wages and number of days worked in a job per year. We utilize this information and distinguish in one specification Yugoslavian workers by their wage growth and group them by whether their wage growth was above or below the median of workers considered in the treatment variable. To so we use all workers with positive wage information in full-time employment and compute the compound average growth rate for the first and last observed wage. We also group them not only by their own wages but also by the wage levels of their employers. Therefore

¹⁷⁴Because we use a relatively short period to construct the treatment variable, the exact definition of the main job has no major influence. About 92% of the workers in our treatment variable do not change the industry and 89% do not change their occupation.

¹⁷⁵Namely managers, skilled commercial and administration, professions, semi-professions, skilled services, engineers, technicians, and skilled manual occupations.

we group workers by whether they worked (or not) for one of the employers with a median establishment wage in the top quartile of the industry. For this we computed the quartiles of establishment wages by industry, based on IAB Established History Panel (BHP). The BHP comprises the universe of German establishments with at least one employee subject to social security contributions. Our data and the BHP can be linked via a common identifier (for more information on the BHP see Schmucker et al., 2018).

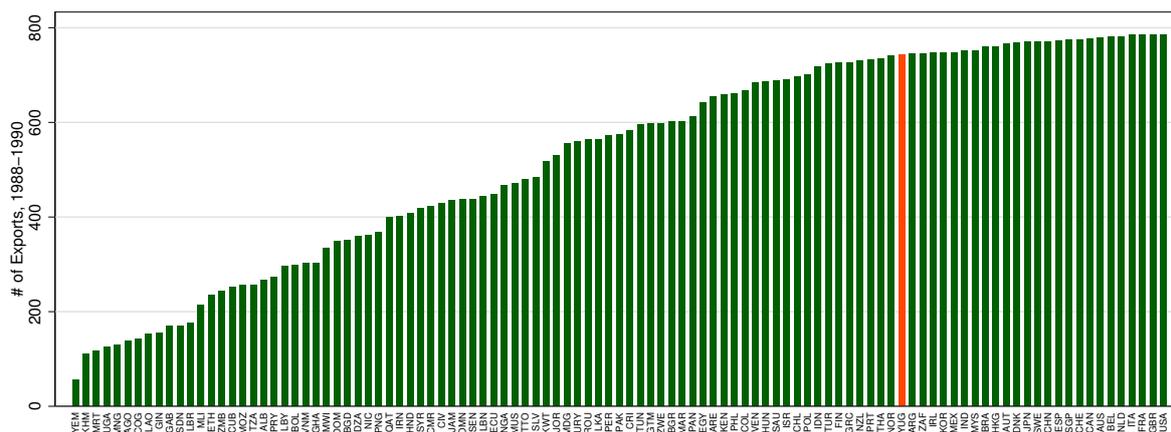
4.10.3 Zeroes in the Data

There are 38 products that Yugoslavia does not export in either the pre-treatment period (1988-1990) or the post-treatment period (2005-2007). These products are excluded from our model when we examine log exports, but included in two other specifications. Including these products has a large impact on the magnitude of our estimated treatment effect, doubling the size of the coefficient in our instrumental variable specification (see Table 32 in main text).

For this reason, we look more closely at the prevalence of zero export products in this appendix. If the zeros mostly occur in the pre-treatment period, we might conclude that returning migrants launched new industries in Yugoslavia, which would explain the increased size of the treatment effect. However, this is not the case. Between the pre- and post-periods, 12 product lines were opened and 16 product lines were closed, and the product lines that were closed are much larger than those that opened. Three of the closed product lines are especially large, with exports over \$20 million in the pre-period but nothing in the post period. All three of these products are liquid fuels. Our results are robust to the exclusion of these three fuels, and we find that those products alone do not cause the increase in the size of the estimated treatment effect.

Given that Yugoslavia does export 770 of 786 products in the pre-treatment period, we examine the total number of products exported by other countries in that period to ascertain whether Yugoslavia is unusual in having so many export lines. We find that it is in fact not uncommon for countries to export so many products, and several developing markets a comparable levels of GDP per capita have a greater number of exports. This is shown in Figure 36.

Figure 36: Exploration of zeros in our data



The figure compare the number of products exported to the rest of the world with reported export value above zero in the baseline period. The figure shows that Yugoslavia is no outlier in terms of the number of products exported (or, alternatively, the number of products with export value equal to zero).

We also consider the possibility that, though Yugoslavia exports many products, most of these export lines are small and insignificant. If this were true, our use of product fixed effects means that our results

could be produced by sectors that are largely unimportant to the former Yugoslavian economies today. We therefore ran our main results excluding products with fewer than \$25,000 in exports in the pre-treatment period. Our results hold using this sub-sample, and we see that just 81 of 786 products have exports of less than \$25,000. We therefore conclude that the change in our estimated treatment effect changes when we add products with zero exports does not reflect a larger pattern of zeros driving our results.

4.10.4 Including Exports to Germany

Our main estimations purposely excludes exports to Germany from the dependent variable, to avoid our results being driven by a reduction of transaction costs following return migration. Table 41 presents results with total exports from Yugoslavian countries to the rest of the world, including to Germany. As expected, our point estimates are larger than in the main results.

Table 41: DID, including exports to Germany

Dependent variable: $exports_{p,t}$	OLS			2SLS		
	lnexp	lnexpplus1	asinhexp	lnexp	lnexpplus1	asinhexp
treat2000 \times after2005	0.0837 (0.038)**	0.1358 (0.063)**	0.1376 (0.066)**	0.1133 (0.051)**	0.2395 (0.086)***	0.2449 (0.089)***
germanexpshare1990 \times after2005	6.3680 (4.325)	-0.1571 (6.815)	-0.4415 (7.001)	3.3411 (5.054)	-11.3455 (12.364)	-12.0176 (12.800)
N	1496	1572	1572	1496	1572	1572
r2	0.86	0.81	0.81	0.86	0.81	0.81
KP F Stat				661.95	723.55	723.55

This table shows result of the estimation for specification (36) using different monotonic transformations for $exports_{p,t}$ in each column. The estimation uses average exports for years 1988 to 1990 in the initial period and average exports for years 2005 to 2007 in the end period. The dependent variable includes exports to Germany. The first three columns report results from an OLS estimation, while the last three columns report results from a 2SLS estimation. All columns include product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.10.5 Additional Controls

Some endogeneity concerns might remain given the lack of use of a measure of FDI stocks in Yugoslavia which include not only Germany but most other countries. The reason for the concern is that the inflow of workers with skills relevant to a particular industry might trigger FDI into the country, from several sources, which can be the explanatory source of the rise of exports that we document. In order to deal with that, we gather data on aggregate FDI stocks in Yugoslavian countries by 2-digit SITC products in 2005. Given the lack of data of FDI stocks in Yugoslavia disaggregated by product before its dissolution, we assume the stock was zero in the “before” period of 1990. The data comes from The Vienna Institute for International Economic Studies (wiiw). Table 42 shows our main results are robust to the inclusion of the global FDI stock in Yugoslavia instead of the FDI from Germany only.

Table 42: DID, controlling for global FDI

Dependent variable: $exports_{p,t}$	OLS			2SLS		
	lnexp	lnexpplus1	asinhexp	lnexp	lnexpplus1	asinhexp
treat2000 \times after2005	0.0730 (0.038)*	0.1245 (0.064)*	0.1262 (0.066)*	0.1278 (0.050)**	0.2591 (0.084)***	0.2646 (0.087)***
germanexpshare1990 \times after2005	9.3848 (4.420)**	3.7292 (5.901)	3.4709 (6.079)	3.6356 (4.999)	-11.0989 (12.322)	-11.7700 (12.760)
lnglobalfdi	-0.1067 (0.050)**	-0.1371 (0.084)	-0.1380 (0.087)	-0.0988 (0.051)*	-0.1214 (0.084)	-0.1219 (0.087)
N	1496	1572	1572	1496	1572	1572
r2	0.86	0.82	0.81	0.86	0.81	0.81
KP F Stat				674.19	737.95	737.95

This table shows result of the estimation for specification (36) using different monotonic transformations for $exports_{p,t}$ in each column. The estimation uses average exports for years 1988 to 1990 in the initial period and average exports for years 2005 to 2007 in the end period. It includes as a control the global FDI stock in Yugoslavia by industry, as opposed to the Germany FDI stock only. The first three columns report results from an OLS estimation, while the last three columns report results from a 2SLS estimation. All columns include product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In addition to this, the results presented in the paper raise some concerns regarding the negative sign of the estimates of the partial correlation of FDI stocks and export levels. If anything, we would expect this control to have a positive sign. To explore what is that drives this unexpected relationship we reestimate a variation of specification (36) that only includes the FDI variable on the right hand side. That is, we are analyzing the partial correlation between exports and FDI in our setting. The results are presented in Table 43. Columns 1-3 uses both product and year fixed effects, while columns 4-6 only uses year fixed effects. We can see that when excluding the product fixed effects the partial correlation between exports and FDI is estimated to be a positive one, as expected. This occurs, plausibly, because introducing the product fixed effects leave very little variation to be use in the estimation of the FDI coefficient, particularly because the FDI figures are defined at the 2 digit level, and the fixed effects at the 4-digit level. All in all, we find that when excluding the product fixed effects, products that have

more FDI during that period explain larger exports, as it is to be expected.

Table 43: Exports vs. FDI

Dependent variable: $exports_{p,t}$						
	lnexp	lnexpplus1	asinhexp	lnexp	lnexpplus1	asinhexp
lnfdi	-0.1518 (0.066)**	-0.2594 (0.125)**	-0.2640 (0.130)**	0.1421 (0.062)**	0.2519 (0.090)***	0.2573 (0.092)***
N	1496	1572	1572	1524	1572	1572
r2	0.86	0.82	0.81	0.03	0.02	0.02
Product FE	Y	Y	Y	N	N	N

This table shows result of the estimation for specification (36) that only includes FDI stock as the right hand side variable, using different monotonic transformations for $exports_{p,t}$ in each column. The first 3 columns include product fixed effects and the following 3 columns do not include those product fixed effects. All columns include year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Finally, we also test for and rule out an additional hypotheses: returning refugees fostering imports of intermediate inputs needed for the production of the exported good. To do so we reestimate our main specification adding as a control the total imports from Germany of intermediate goods for each product based on input-output relationships. This exercise goes by incorporating an input-output matrix between SITC industries. We construct this matrix based on the US historical benchmark input output tables put together by the Bureau of Economic Activity for year 2002 based on the NAICS classification.¹⁷⁶ We then match the input-output coefficients to SITC 4-digits using the same procedure as Cuñat and Melitz (2012) described in their footnote 24 and their subsequent documentation. Note that both the input-output coefficients as well as the concordances is based on US data. While not ideal, we believe this is a good approximation for our purposes.

We then incorporate in the estimation as a control, for each product p , the weighted sum of imports from Germany based on the input-output coefficient (e.g., US dollars of each input needed to produce 1 US\$ of output), both in the "before" and the "after" period (using the inverse asymptotic sine transformation). This estimation is presented in Table 44, and they are robust to our main results.

¹⁷⁶The data was downloaded from <https://www.bea.gov/industry/historical-benchmark-input-output-tables> in May of 2019.

Table 44: DID, controlling for imports of intermediate goods

Dependent variable: $exports_{p,t}$	OLS			2SLS		
	lnexp	lnexpplus1	asinhexp	lnexp	lnexpplus1	asinhexp
treat2000 \times after2005	0.0751 (0.037)**	0.1472 (0.076)*	0.1500 (0.079)*	0.1101 (0.050)**	0.2416 (0.101)**	0.2468 (0.106)**
germanexpshare1990 \times after2005	0.1712 (6.524)	-10.4364 (12.188)	-10.9717 (12.573)	-3.3168 (8.070)	-20.6303 (17.331)	-21.4251 (17.885)
lnioimpdeu	0.0981 (0.150)	0.0746 (0.309)	0.0760 (0.323)	0.1035 (0.150)	0.0756 (0.310)	0.0770 (0.323)
N	1428	1572	1572	1428	1572	1572
r2	0.86	0.79	0.79	0.86	0.79	0.79
KP F Stat				605.00	711.53	711.53

This table shows result of the estimation for specification (36) using different monotonic transformations for $exports_{p,t}$ in each column. The estimation uses average exports for years 1988 to 1990 in the initial period and average exports for years 2005 to 2007 in the end period. It includes as a control the total imports from Germany in each period of intermediate goods used in the production of the product under consideration (weighted by 2002 input-output coefficients based on data from the US Bureau of Economic Activity). The first three columns report results from an OLS estimation, while the last three columns report results from a 2SLS estimation. All columns include product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.10.6 Further Tests on the Exogeneity of Exit From the Labor Force

Figure 30 presents a very convincing picture in terms of the exogeneity of exit from the labor force. Yet, we still see some important outliers in that graph as measured by their distance to the 45 degree line.

This, and all the other smaller deviations, could happen in part because we know that not all of those entering the labor force between 1991 and 1995 are in fact *Duldung* holders given that some of them stayed on beyond the year 2000. Yet, given this fact, the correlation between the two measures is quite high.

But yet again, the question is whether the exit from the labor force is exogenous with respect to the dynamics of exports at the product level back in Yugoslavia. In order to find whether the outliers are an issue to be concern about, we first compute a measure of the deviations from the 45 degree line which is the difference between the horizontal axis and the vertical axis. If this measure is above 0 for a given industry, for instance, it means the proportion of those Yugoslavians exiting from that industry was higher than the proportion of those who arrived to it.

We correlate this measure with the growth rate of Yugoslavian exports by industry between 1990 and 2005. Using three different measures of export growth that subtract the 1990 from the 2005 value after transforming export values in both periods using $\log(x)$, $\log(x + 1)$ and $\operatorname{asinh}(x)$, the correlations result in -0.0295, -0.0138, -0.0133, respectively, all statistically insignificant.

4.10.7 Robustness: German WZ 93 3-digit Industries

Table 45 replicates our main results based on the German WZ 93 3-digit industry classification, as opposed to SITC 4-digits. Naturally, using a higher-level of aggregation results in many fewer observations (about 200 as opposed to 1500) and with that much less variation both on the dependent and independent variables. In this case, with barely enough variation to exploit, we still find results that are qualitatively similar (besides in Columns 1 and 4) and that are statistically significant in columns 5 and 6 that correspond to the 2SLS estimator. Note that the reported first-stage F statistics of the 2SLS estimators are strong, but the number is not as high as the main results.

Table 45: DID, WZ 93 3-digit aggregation

Dependent variable: $exports_{p,t}$	OLS			2SLS		
	lnexp	lnexpplus1	asinhexp	lnexp	lnexpplus1	asinhexp
treat2000 × after2005	-0.1219 (0.084)	0.0636 (0.041)	0.0636 (0.041)	-0.0910 (0.119)	0.1343 (0.040)***	0.1343 (0.040)***
germanexpshare1990 × after2005	3.1768 (2.847)	1.5130 (2.892)	1.5130 (2.892)	2.0230 (3.882)	-2.9765 (4.530)	-2.9765 (4.530)
N	226	282	282	226	282	282
r2	0.93	1.00	1.00	0.93	0.99	1.00
KP F Stat				155.46	847.68	847.68

This table shows result of the estimation for specification (36) based using data disaggregated based on German WZ 93 3-digit industries as opposed to SITC 4-digit ones. The results are presented using different monotonic transformations for $exports_{p,t}$ in each column. The estimation uses average exports for years 1988 to 1990 in the "before" period and average exports for years 2005 to 2007 in the "after" period. The first three columns report results from an OLS estimation, while the last three columns report results from a 2SLS estimation. All columns include product fixed effects and year fixed effects. Standard errors clustered at the W93 industry level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 46 replicates our main results using SITC 4-digit disaggregation, but the standard errors are clustered using groups that correspond to the German WZ 93 3-digit industry classification. We do this because both our treatment and instrumental variables are originally defined at that level of aggregation, before we disaggregate them in SITC 4-digit industries. The results are robust to the use of these higher-level of aggregation when defining clusters to estimating the standard errors.

Table 46: DID, clustering by W93 industries

Dependent variable: $exports_{p,t}$	OLS			2SLS		
	lnexp	lnexpplus1	asinhexp	lnexp	lnexpplus1	asinhexp
	treat2000 \times after2005	0.0837 (0.052)	0.1358 (0.075)*	0.1376 (0.078)*	0.1133 (0.057)*	0.2395 (0.097)**
germanexpshare1990 \times after2005	6.3680 (5.022)	-0.1571 (6.793)	-0.4415 (6.959)	3.3411 (6.065)	-11.3455 (13.204)	-12.0176 (13.640)
N	1496	1572	1572	1496	1572	1572
r2	0.86	0.81	0.81	0.86	0.81	0.81
KP F Stat				174.03	199.68	199.68

This table shows result of the estimation for specification (36) using different monotonic transformations for $exports_{p,t}$ in each column. The estimation uses average exports for years 1988 to 1990 in the "before" period and average exports for years 2005 to 2007 in the "after" period. The first three columns report results from an OLS estimation, while the last three columns report results from a 2SLS estimation. All columns include product fixed effects and year fixed effects. Standard errors clustered at the W93 industry level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.10.8 Other Robustness Checks

Table 47 reports 2SLS estimation for specification (36) using alternatives treatments. The first three columns use the difference in Yugoslavian workers per industry between 2005 and 1995, and the following three columns simply uses the baseline stock of Yugoslavian workers per industry in 1995.

Table 47: DID, different treatments

Dependent variable: $exports_{p,t}$						
	lnexp	lnexp	lnexpplus1	lnexpplus1	asinhexp	asinhexp
treat2005 \times after2005	0.1708 (0.050)***		0.3170 (0.079)***		0.3235 (0.081)***	
treat1995level \times after2005		0.1388 (0.040)***		0.2557 (0.063)***		0.2609 (0.065)***
germanexpshare1990 \times after2005	1.3106 (5.667)	1.0313 (5.710)	-12.0252 (10.353)	-12.4511 (10.041)	-12.6154 (10.679)	-13.0500 (10.356)
N	1520	1520	1572	1572	1572	1572
r2	0.86	0.86	0.80	0.80	0.80	0.80
KP F Stat	700.62	865.62	764.86	960.95	764.86	960.95

This table shows result of the 2SLS estimation for specification (36) using different monotonic transformations for $exports_{p,t}$ in each column. The estimation uses two other different definitions of treatment: (i) return migrants between 1995 and 2005, and (ii) the stock of migrant workers in 1995. The estimation uses exports between 1990 and 2005. All columns include product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 48 reports 2SLS estimation for specification (36) using only observations with non-zero exports for the two monotonic transformations of the dependent variable $\log(exports_{c,p,t}+1)$ and $asinh(exports_{c,p,t})$.

Table 48: DID, no zeros

Dependent variable: $exports_{p,t}$				
	OLS		2SLS	
	lnexpplus1	asinhexp	lnexpplus1	asinhexp
treat2000 \times after2005	0.0837 (0.038)**	0.0837 (0.038)**	0.1133 (0.051)**	0.1133 (0.051)**
germanexpshare1990 \times after2005	6.3686 (4.325)	6.3680 (4.325)	3.3412 (5.054)	3.3411 (5.054)
N	1496	1496	1496	1496
r2	0.86	0.86	0.86	0.86
KP F Stat			661.95	661.95

This table shows result of the 2SLS estimation for specification (36) using different monotonic transformations for $exports_{p,t}$ in each column, excluding observations for which there were zero exports in either period. The estimation uses years 1990 and 2005. All columns include product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.10.9 Excluding Slovenia and Macedonia from the Sample

Table 49 replicates the results of Table 34 excluding Slovenia from the sample. Slovenia was the first Yugoslavian republic to secede and did not suffer from a long war nor a massive exile of its inhabitants to other locations. Our results are robust to its exclusion from the left hand side variable.

Table 49: DID, excl. Slovenia

Dependent variable: $exports_{p,t}$	OLS			2SLS		
	lnexp	lnexpplus1	asinhexp	lnexp	lnexpplus1	asinhexp
	treat2000 \times after2005	0.0837 (0.038)**	0.1358 (0.063)**	0.1376 (0.066)**	0.1133 (0.051)**	0.2395 (0.086)**
germanexpshare1990 \times after2005	6.3680 (4.325)	-0.1571 (6.815)	-0.4415 (7.001)	3.3411 (5.054)	-11.3455 (12.364)	-12.0176 (12.800)
N	1496	1572	1572	1496	1572	1572
r2	0.86	0.81	0.81	0.86	0.81	0.81
KP F Stat				661.95	723.55	723.55

This table shows result of the estimation for specification (36) using different monotonic transformations for $exports_{p,t}$ in each column, excluding exports from Slovenia as one of the former Yugoslavian republics post 1992. The estimation uses years 1995 and 2005. All columns include product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 50 repeats the same exercise and excludes Macedonia. Following Slovenia and Croatia, Macedonia held a referendum and declared its independence in late 1991. Unlike others, Macedonia managed to obtain its independence without going through an armed conflict. This is why, no war refugees from Macedonia went to Germany. Our results are robust to its exclusion from the left hand side variable.

Table 50: DID, excl. Macedonia

Dependent variable: $exports_{p,t}$	OLS			2SLS		
	lnexp	lnexpplus1	asinhexp	lnexp	lnexpplus1	asinhexp
	treat2000 \times after2005	0.0837 (0.038)**	0.1358 (0.063)**	0.1376 (0.066)**	0.1133 (0.051)**	0.2395 (0.086)**
germanexpshare1990 \times after2005	6.3680 (4.325)	-0.1571 (6.815)	-0.4415 (7.001)	3.3411 (5.054)	-11.3455 (12.364)	-12.0176 (12.800)
N	1496	1572	1572	1496	1572	1572
r2	0.86	0.81	0.81	0.86	0.81	0.81
KP F Stat				661.95	723.55	723.55

This table shows result of the estimation for specification (36) using different monotonic transformations for $exports_{p,t}$ in each column, excluding exports from Macedonia as one of the former Yugoslavian republics post 1992. The estimation uses years 1995 and 2005. All columns include product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.10.10 Occupations by Characteristics

Table 51 presents the list of all the occupations in the dataset, with their respective frequency, and associated characteristics.

Table 51: Occupations List

Occupation	Occurrences	Tasks	Prof. skills	Share supervisor
323 Metal workers (no further specification)	968	2 manual	1 unskilled	.21722362
51 Gardeners, garden workers	858	2 manual	1 unskilled	.18927162
531 Assistants (no further specification)	787		1 unskilled	.11138389
522 Packagers, goods receivers, despatchers	719	2 manual	1 unskilled	.09713266
151 Plastics processors	698	2 manual	1 unskilled	.27068706
181 Wood preparers	602	2 manual	1 unskilled	.09192798
391 Bakery goods makers	555	2 manual	2 skilled	.36883311
744 Stores, transport workers	543	2 manual	1 unskilled	.10129447
501 Carpenters	498	2 manual	2 skilled	.23684861
62 Forest workers, forest cultivators	483	2 manual	1 unskilled	.20126734
402 Meat, sausage goods makers	442	2 manual	1 unskilled	0
271 Building fitters	396	2 manual	2 skilled	.15821244
470 Building labourer, general	386			
933 Household cleaners	376	2 manual	1 unskilled	.07903877
322 Other assemblers	312	2 manual	1 unskilled	.04823485
411 Cooks	266	2 manual	2 skilled	.28007612
682 Salespersons	263	1 analytic	1 unskilled	.1624328
101 Stone preparers	257	2 manual	1 unskilled	.23494546
141 Chemical plant operatives	257	1 analytic	1 unskilled	.28136134
321 Electrical appliance, electrical parts assemblers	250	2 manual	1 unskilled	.14907955
275 Steel structure fitters, metal shipbuilders	250	2 manual	2 skilled	.19518355
112 Shaped brick, concrete block makers	244	2 manual	1 unskilled	.22286043
401 Butchers	239	2 manual	2 skilled	.29524547
442 Concrete workers	232	2 manual	1 unskilled	.41252334
163 Book binding occupations	210	2 manual	2 skilled	.13312343
273 Engine fitters	186	1 analytic	2 skilled	.2418478
521 Goods examiners, sorters, n.e.c.	183	1 analytic	1 unskilled	.17720443
241 Welders, oxy-acetylene cutters	173	2 manual	1 unskilled	.16681339
412 Ready-to-serve meals, fruit, vegetable preservers, preparers	169	2 manual	1 unskilled	0
714 Motor vehicle drivers	168	2 manual	1 unskilled	.07980934
311 Electrical fitters, mechanics	151	1 analytic	2 skilled	.31535207
741 Warehouse managers, warehousemen	146	1 analytic	1 unskilled	.36790809
352 Clothing sewers	141	2 manual	1 unskilled	0
392 Confectioners (pastry)	136	2 manual	2 skilled	.23874688
41 Land workers	135	2 manual	1 unskilled	.03541072
431 Milk, fat processing operatives	119	2 manual	1 unskilled	.18633992
211 Sheet metal pressers, drawers, stampers	111	2 manual	1 unskilled	0
270 Locksmiths, not specified	108	2 manual	2 skilled	.27546819
433 Sugar, sweets, ice-cream makers	107	2 manual	1 unskilled	.55506282
441 Bricklayers	105	2 manual	2 skilled	.39556132
177 Printer's assistants	102	2 manual	1 unskilled	.19699555
432 Flour, food processors	98	2 manual	1 unskilled	.27100673
143 Rubber makers, processors	95	2 manual	1 unskilled	.24792283
482 Insulators, proofers	85	2 manual	1 unskilled	.35678298
221 Turners	84	1 analytic	2 skilled	.23881324
512 Goods painters, lacquerers	83	2 manual	1 unskilled	.25905095
164 Other paper products makers	79	2 manual	1 unskilled	0
261 Sheet metal workers	79	2 manual	2 skilled	.34136059
451 Carpenters	76	2 manual	2 skilled	.32004301

This table presents the first part of the list of all the occupations in the dataset, with their respective frequency and associated characteristics.

Table 52: Occupations List (cont.)

Occupation	Occurrences	Tasks	Prof. skills	Share supervisor
111 Stoneware, earthenware makers	74		1 unskilled	0
212 Wire moulders, processors	72	2 manual	1 unskilled	.3442623
303 Dental technicians	71	2 manual	2 skilled	.19073161
272 Sheet metal, plastics fitters	68		2 skilled	0
472 Other building labourers, building assistants, n.e.c.	66	2 manual	1 unskilled	.12660338
742 Transportation equipment drivers	66	2 manual	1 unskilled	.0789034
781 Office specialists	66	1 analytic	2 skilled	.26168916
162 Packaging makers	60	2 manual	1 unskilled	.04646367
135 Glass processors, glass finishers	60	2 manual	1 unskilled	.14649977
121 Ceramics workers	59	2 manual	1 unskilled	.05361638
274 Plant fitters, maintenance fitters	57	2 manual	2 skilled	.23151613
161 Paper, cellulose makers	54	2 manual	1 unskilled	.49392581
71 Miners	53	2 manual	1 unskilled	.1553002
251 Steel smiths	50	2 manual	2 skilled	.04485785
263 Pipe, tubing fitters	49	2 manual	1 unskilled	.3949083
262 Plumbers	49	2 manual	2 skilled	.26210474
356 Sewers, n.e.c.	48	2 manual	1 unskilled	.04908014
492 Upholsterers, mattress makers	48	2 manual	2 skilled	.12503124
923 Other housekeeping attendants	47	2 manual	1 unskilled	.08383468
376 Leather clothing makers and other leather processing operatives	45	2 manual	1 unskilled	
373 Footwear makers	44	2 manual	1 unskilled	0
937 Machinery, container cleaners and related occupations	44	2 manual	1 unskilled	.05061111
81 Stone crushers	43		1 unskilled	0
313 Electric motor, transformer fitters	39	1 analytic	1 unskilled	.13471446
485 Glaziers	39	2 manual	2 skilled	.1977815
403 Fish processing operatives	36	2 manual	1 unskilled	.5
371 Leather makers, catgut string makers	35	2 manual	1 unskilled	0
225 Metal grinders	35	2 manual	1 unskilled	.28720212
284 Precision mechanics	35	1 analytic	2 skilled	.16234579
285 Other mechanics	34	2 manual	2 skilled	.2512635
462 Road makers	33	2 manual	1 unskilled	.26799082
291 Toolmakers	32	2 manual	2 skilled	.21778998
224 Borers	31	2 manual	1 unskilled	0
931 Laundry workers, pressers	31	2 manual	1 unskilled	.1339676
222 Drillers	30	1 analytic	1 unskilled	.15145272
82 Earth, gravel, sand quarriers	29		1 unskilled	
282 Agricultural machinery repairers	29	2 manual	2 skilled	.30161076
466 Other civil engineering workers	28	2 manual	1 unskilled	.42729718
314 Electrical appliance fitters	28	1 analytic	2 skilled	.26680461
44 Animal keepers and related occupations	25	2 manual	1 unskilled	.35544285
281 Motor vehicle repairers	25	1 analytic	2 skilled	.34470057
234 Galvanisers, metal colourers	24	2 manual	1 unskilled	.14969613
213 Other metal moulders (non-cutting deformation)	24	2 manual	1 unskilled	0
203 Semi-finished product fettlers and other mould casting occupations	24	2 manual	1 unskilled	.1735251
331 Spinners, fibre preparers	24	2 manual	2 skilled	.48452174
935 Street cleaners, refuse disposers	22	2 manual	1 unskilled	.30638207
342 Weavers	21	2 manual	1 unskilled	0
353 Laundry cutters, sewers	21	2 manual	1 unskilled	0
344 Machined goods makers	21	2 manual	1 unskilled	0
423 Other beverage makers, tasters	20	2 manual	2 skilled	.64293598
784 Office auxiliary workers	20	1 analytic	1 unskilled	.09301868

This table presents the second part of the list of all the occupations in the dataset, with their respective frequency and associated characteristics.

4.10.11 Estimations Using Treatments by Educational Attainment and Occupations Characteristics

Table 40 in the main body of the paper summarized the results exploiting heterogeneity of the treatment in terms of the skills and occupation characteristics of the workers. Tables 53 to 58 below present results for each estimation separately.

Table 53: Difference-in-differences (OLS), workers weighted by wage

Dependent variable: $exports_{p,t}$	(1)	(2)	(3)	(4)	(5)	(6)
	lnexp	lnexp	lnexpplus1	lnexpplus1	asinhexp	asinhexp
treatwageKsm1 \times after2005	0.1479 (0.070)**	0.1479 (0.070)**	0.2352 (0.100)**	0.2352 (0.100)**	0.2382 (0.103)**	0.2382 (0.103)**
N	1496	1496	1572	1572	1572	1572
Adj R2	0.86	0.86	0.81	0.81	0.81	0.81

This table shows result of the estimation for specification (36) using a treatment constructed by weighting each worker by his/her wage (in thousand euros). The table presents OLS estimations. Each group of results uses different monotonic transformations for $exports_{p,t}$ in different columns. All columns include FDI as control, as well as product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

** $p < 0.10$, * $p < 0.05$, *** $p < 0.01$

Table 54: Difference-in-differences (OLS), workers' education

Dependent variable: $exports_{p,t}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	lnexp	lnexp	lnexp	lnexpplus1	lnexpplus1	lnexpplus1	asinhexp	asinhexp	asinhexp
treatedulow \times after2005	0.0762 (0.041)*		-0.2794 (0.121)**	0.1323 (0.065)**		-0.1958 (0.140)	0.1343 (0.067)**		-0.1923 (0.142)
treateduhig \times after2005		0.1245 (0.043)***	0.4128 (0.128)***		0.1780 (0.068)***	0.3810 (0.140)***		0.1798 (0.071)**	0.3791 (0.142)***
N	1496	1496	1496	1572	1572	1572	1572	1572	1572
Adj R2	0.86	0.86	0.86	0.81	0.81	0.82	0.81	0.81	0.81

This table shows result of the estimation for specification (36) using treatments constructed by aggregating workers by groups based on their educational level. The table presents OLS estimations. Each group of results uses different monotonic transformations for $exports_{p,t}$ in different columns. All columns include FDI as control, as well as product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 55: Difference-in-differences (OLS), workers' occupation tasks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable: $exports_{p,t}$	lnexp	lnexp	lnexp	lnexpplus1	lnexpplus1	lnexpplus1	asinhexp	asinhexp	asinhexp
treattaskmanual \times after2005	0.0748 (0.038)*		0.0222 (0.054)	0.1269 (0.061)**		0.0705 (0.074)	0.1287 (0.063)**		0.0721 (0.076)
treattaskanalytical \times after2005		0.1575 (0.055)***	0.1300 (0.077)*		0.2289 (0.081)***	0.1397 (0.080)*		0.2315 (0.083)***	0.1402 (0.080)*
N	1496	1496	1496	1572	1572	1572	1572	1572	1572
Adj R2	0.86	0.86	0.86	0.81	0.81	0.81	0.81	0.81	0.81

This table shows result of the estimation for specification (36) using treatments constructed by aggregating workers by groups based on the intensity of tasks linked to their occupations (manual vs. analytical). The table presents OLS estimations. Each group of results uses different monotonic transformations for $exports_{p,t}$ in different columns. All columns include FDI as control, as well as product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 56: Difference-in-differences (OLS), workers' occupation skill level

Dependent variable: $exports_{p,t}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	lnexp	lnexp	lnexp	lnexpplus1	lnexpplus1	lnexpplus1	lnexpplus1	lnexpplus1	lnexpplus1
treatbf2lowskill \times after2005	0.0778 (0.040)*		0.0071 (0.064)	0.1317 (0.066)**		0.0592 (0.094)	0.1335 (0.068)**		0.0609 (0.097)
treatbf2higskill \times after2005		0.1272 (0.044)***	0.1199 (0.071)*		0.1864 (0.064)***	0.1242 (0.080)		0.1885 (0.065)***	0.1245 (0.081)
N	1496	1496	1496	1572	1572	1572	1572	1572	1572
Adj R2	0.86	0.86	0.86	0.81	0.81	0.81	0.81	0.81	0.81

This table shows result of the estimation for specification (36) using treatments constructed by aggregating workers by groups based on the skill level of their occupation. The table presents OLS estimations. Each group of results uses different monotonic transformations for $exports_{p,t}$ in different columns. All columns include FDI as control, as well as product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

Table 57: Difference-in-differences (OLS), workers' occupation supervision intensity

Dependent variable: $exports_{p,t}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	lnexp	lnexp	lnexp	lnexpplus1	lnexpplus1	lnexpplus1	asinhexp	asinhexp	asinhexp
treatsvct0 × after2005	0.0873 (0.039)**		-0.0554 (0.108)	0.1398 (0.065)**		0.0200 (0.175)	0.1416 (0.067)**		0.0218 (0.182)
treatsvct1 × after2005		0.1435 (0.050)***	0.2174 (0.137)		0.2118 (0.075)***	0.1848 (0.197)		0.2143 (0.077)***	0.1849 (0.202)
N	1496	1496	1496	1572	1572	1572	1572	1572	1572
Adj R2	0.86	0.86	0.86	0.81	0.81	0.81	0.81	0.81	0.81

This table shows result of the estimation for specification (36) using treatments constructed by aggregating workers by groups based on the supervision intensity of their occupation. The table presents OLS estimations. Each group of results uses different monotonic transformations for $exports_{p,t}$ in different columns. All columns include FDI as control, as well as product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 58: Difference-in-differences (OLS), workers' wage growth

Dependent variable: $exports_{p,t}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	lnexp	lnexp	lnexp	lnexpplus1	lnexpplus1	lnexpplus1	asinhexp	asinhexp	asinhexp
treatwgrcagramd0 \times after2005	0.0867 (0.045)*		-0.0405 (0.212)	0.1435 (0.069)**		-0.0358 (0.225)	0.1455 (0.071)**		-0.0353 (0.226)
treatwgrcagramd1 \times after2005		0.0912 (0.044)**	0.1311 (0.210)		0.1493 (0.070)**	0.1845 (0.231)		0.1512 (0.072)**	0.1860 (0.232)
N	1496	1496	1496	1572	1572	1572	1572	1572	1572
Adj R2	0.86	0.86	0.86	0.81	0.81	0.81	0.81	0.81	0.81

This table shows result of the estimation for specification (36) using treatments constructed by aggregating workers by groups based on the wage growth during their stay. The table presents OLS estimations. Each group of results uses different monotonic transformations for $exports_{p,t}$ in different columns. All columns include FDI as control, as well as product fixed effects and year fixed effects. Standard errors clustered at the product level presented in parenthesis. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

4.10.12 Expanding to All Countries: External Validation

After having established the link between migration and comparative advantage, we turn to study the same phenomenon in a multi-country and multi-period setting. In this setting our focus is not on the identification, but rather on externally validating the results, while exploiting a much larger variation allowing us to study differential effects based on the characteristics of the migrants. That is, we expand our difference-in-difference strategy to all countries in the original dataset using as treatment the presence and sizes of their diasporas in Germany working in different 4-digit products.

Our prior for this exercise is that if knowledge diffusion is the mechanisms through which migration explains productivity shifts seen as changes in the comparative advantage of nations, this effect should be stronger among migrants that are skilled and/or work in occupations that are more cognitive and analytical in nature. This is what we explore in this section.

Empirical strategy and summary statistics In this section we adapt our difference-in-difference specification to a multi-country multi-period setting. To do that, we follow Besley and Burgess (2004) and estimate the following specification:¹⁷⁷

$$exports_{c,p,t} = \beta^{DID} migrants_{c,p,t-10} + \beta^{ge} globalexports_{p,t} + \eta_{c,p} + \alpha_{c,t} + \varepsilon_{c,p,t} \quad (37)$$

Our dependent variable, $exports_{c,p,t}$, is defined as total export value of product p during year t from country c to the rest of the world, excluding Germany in order to rule out that our results are driven by lower costs to export due to migrant networks. Similarly to the previous section, we present our results for different monotonic transformations of the dependent variable. Our variable of interest, the treatment, in this case is $migrants_{c,p,t-10}$, which is the stock of migrants from country c at time $t - 10$ (e.g., we allow for a 10-year lag for the treatment to "kick in") working in product p in the German labor force. We also include a series of fixed effects, crucial for the estimation. Since we have expanded

¹⁷⁷Both specifications are equivalent. To see it, suppose the following two specifications, the first one where the treatment is defined as a difference and the second one where the treatment is defined as a level:

$$\begin{aligned} y_{p,t} &= \beta_1 \Delta migrants_p \times after_t + \delta_t + \eta_p + \varepsilon_{p,t} \\ y_{p,t} &= \beta_2 migrants_{p,t} + \delta_t + \eta_p + \varepsilon_{p,t} \end{aligned}$$

Assume there are only two periods, $t = [0, 1]$. According to the first functional form, we have:

$$\begin{aligned} E(y_{p,t}|t=1) &= \beta_1 \Delta migrants_p + \delta_1 + \eta_p + \varepsilon_{p,1} \\ E(y_{p,t}|t=0) &= \delta_0 + \eta_p + \varepsilon_{p,0} \end{aligned}$$

It is clear that $E(y_{p,t}|t=1) - E(y_{p,t}|t=0) = \beta_1 \Delta migrants_p + (\delta_1 - \delta_0) + (\varepsilon_{p,1} - \varepsilon_{p,0})$. According to the second functional form, we have:

$$\begin{aligned} E(y_{p,t}|t=1) &= \beta_2 migrants_{p,1} + \delta_1 + \eta_p + \varepsilon_{p,1} \\ E(y_{p,t}|t=0) &= \beta_2 migrants_{p,0} + \delta_0 + \eta_p + \varepsilon_{p,0} \end{aligned}$$

Thus, in this case, $E(y_{p,t}|t=1) - E(y_{p,t}|t=0) = \beta_2 (migrants_{p,1} - migrants_{p,0}) + (\delta_1 - \delta_0) + (\varepsilon_{p,1} - \varepsilon_{p,0})$. Since $\Delta migrants_p = migrants_{p,1} - migrants_{p,0}$ it follows that $\beta_1 = \beta_2$.

the dimension of our dataset to include countries our unit of analysis becomes now a country-product pair. Thus, we include $\eta_{c,p}$ which is a country-by-product fixed effects, to allow each country-product to have a different intercept and also, in the difference-in-difference setting, allows us to exploit within country-product variation. We also include $\alpha_{c,t}$, a country-by-year fixed effect, which controls for changes at the country level that could explain changes in exports: income, population, institutions, etc. We also include $globalexports_{p,t}$, which in measures the total export value of product p by all countries during year t , to control for total global demand, and as a proxy for the introduction of a technology that explains a global increase in the exports of product p .¹⁷⁸ All of the continuous right hand side variables are monotonically transformed using the inverse hyperbolic sine. Our estimations cluster standard errors at the country-product level (Besley and Burgess, 2004; Bertrand et al., 2004).

As mentioned earlier, the sample for this estimation includes 124 countries and 786 products across two periods: 1990 to 2000 and 2000 to 2010. The IAB data allows us to compute the migrant stock by different categories, and we exploit that variation in this setting. Table 59 summarizes the statistics for the main variables used in this analysis. The first three rows summarize the export value averaged across countries, products and years 2000 and 2010, using three different monotonic transformations; note that the number of observations using a simple logarithmic transformation is reduced due to zeros in the sample.

Table 59: Summary statistics, all countries

Variable	N	Mean	sd	Min	Max
Exports (log)	136,684	14.029	3.44	6.9	25.8
Exports (log +1)	179,208	10.700	6.68	0.0	25.8
Exports (asinh)	179,208	11.229	6.95	0.0	26.5
All Migrants	179,208	8.047	127.48	0.0	22,803.5
Skilled	179,208	3.769	63.15	0.0	12,501.7
Unskilled	179,208	4.001	67.83	0.0	11,614.6
White collars	179,208	0.636	7.24	0.0	798.1
Non-white collars	179,208	7.093	121.60	0.0	22,497.6
Analytical & Cognitive tasks	179,208	1.913	26.45	0.0	3,816.8
Manual tasks	179,208	5.531	92.33	0.0	15,918.0
High prob solving	179,208	1.480	21.36	0.0	4,193.2
Low prob solving	179,208	6.273	107.56	0.0	19,721.0

This table presents the sample summary statistics for the variables used to estimate specification (36).

Table 59 shows that the average number of migrant workers in Germany across all countries and 4-digit products for both 1990 and 2000 (e.g., the baseline years) is 8. The number is surprisingly small, but note that this variable has many zeros (in fact, the median value is zero), and there is a mix of countries from many different sizes. This last fact is reflected in both the large standard deviation and upper bound of the variable which reaches a maximum of over 20 thousand workers.

We start by estimating Specification (37) using all workers, without distinction, as the independent

¹⁷⁸Ideally, we would introduce a product-by-year fixed effect but turns out doing so eliminates most of the remaining variation.

variable. The results are presented in Table 60. The elasticity parameter is estimated to be between 0.08 and 0.11, which falls into the lower range of the the results of Section 4.6. In this case, the point estimate when the dependent variable is a simple logarithmic transformation is lower than in the other columns where the monotonic transformation does include the zeros. This suggests, also consistently with the results from Section 4.6, that return migration (this time computed as the difference in the stock) is also explanatory of the extensive margin (e.g., the emergence of new export sectors).

Table 60: DID, all countries

Dependent variable: $exports_{p,t}$			
	(1)	(2)	(3)
	lnexp	lnexpplus1	asinhexp
L10.AllMigrants	0.0846 (0.015)***	0.1252 (0.030)***	0.1232 (0.032)***
Intotalex	0.8935 (0.016)***	0.4403 (0.011)***	0.4595 (0.012)***
N	114288	165060	165060
Adj R2	0.94	0.91	0.90
cpFE	Y	Y	Y

This table shows result of the estimation for specification (36) using different monotonic transformations for $exports_{p,t}$ in each column. The estimation uses years 2000 and 2010 for exports and 1990 and 2000 for migration. All columns include country-by-product fixed effects and country-by-year fixed effects. Standard errors clustered at the country-product level presented in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5 Agglomeration Effects in a Developing Economy: Evidence from Turkey

Abstract

Spatial inequalities in Turkey are a source of considerable policy concern. In this paper, I estimate agglomeration effects for Turkish provinces to shed light on the origins of spatial inequality in productivity and provide evidence from a developing country context which literature needs. I use social security data, an administrative dataset recently made available at the NUTS-3 level, for 81 provinces of Turkey for the period 2008-2013 and carry out a two-step estimation. I use a variety of panel data techniques and historical instruments to deal with estimation concerns. I estimate an elasticity of labor productivity with respect to the density of 0.056-0.06, which is higher than in developed countries and around the levels observed in developing countries. Contrasting the evidence coming from developed countries, I find weak effects for sorting of workers across Turkish provinces based on observable characteristics.

Keywords: local labour markets, spatial wage disparities, developing country, Turkey

JEL Classification Numbers: R12, R23, J31

5.1 Introduction

In the past decade, the empirical literature on agglomeration economies provided robust evidence on the productivity gains associated with larger cities. Despite extensive evidence on these gains in developed countries, little is known about the impact of urbanization in the rest of the world. Addressing the knowledge gap regarding the urban economies in the developing world is important for two reasons. First, the majority of the world's urban population lives in countries that are far poorer than the advanced countries (e.g., US, Europe) where the evidence mainly comes from (Glaeser and Henderson (2017)). In addition to the importance of urban areas being drivers of economic growth in those countries, they also concern the lives of millions of people who reside and work in these places. Second, the models and stylized facts documented for cities in the developed countries may not apply entirely to the developing countries as characteristics and role of cities in the national economy may differ (Chauvin et al. (2017)). For instance, the rapid urbanization observed in the developing countries in the second half of the past century may have generated different benefits and costs compared to the western world where the urbanization rates have been relatively stable.

Turkey is an upper-middle-income developing country that has experienced fast urbanization and has a high rate of the urban population. Since the 1950s, the urban population in Turkey has increased dramatically due to massive rural-migration and high fertility rate. In 2017, 75% of the Turkish population lived in cities, making it a very highly urbanized country (World Bank).¹⁷⁹ In terms of GDP per capita, Turkey has the highest regional disparity among the OECD countries (OECD, 2018), while substantial spatial inequalities also exist in almost every metric (i.e., production, life quality, etc.). In this paper, I study the sources of spatial differences in productivity by focusing on agglomeration effects in Turkey. A better understanding of these effects is crucial for three reasons. First, productivity differences are an important driver of growth in Turkey, and thus, understanding which local factors make a given worker more productive is crucial. Second, understanding the determinant forces would make it possible to formulate policies to reduce regional differences. Third, with a population of 82 million, of which 75% percent living in cities makes agglomeration economies a relevant issue for a vast majority of the population. Despite the increasing urbanization and large spatial wage disparities, the determinants of these differences have not been studied yet. This study aims to fill this gap.

Spatial differences in wages (thus productivity) occur through three main channels.¹⁸⁰ First, differences in wages across areas could be due to differences in skill composition of the workers. Workers with higher skills may sort into more denser areas, if for instance skill-intensive industries are not evenly dis-

¹⁷⁹Turkey went through rapid urbanization starting from the 1950s due to large rural-urban migration. While 34% of the population lived in cities in 1950, 42% in 1975, 53% in 1985, 64.9% in 2000 and it is at 74.5% in 2017 (World Bank). Most of the internal migrants were low-skilled agricultural workers (Kirdar and Saracoglu (2008)). While 62.5% of the Turkish labor force was employed in agriculture in 1980, only 18% remained in this sector as of 2018 (Turkstat).

¹⁸⁰Wages are usually proportional to labor productivity. By using the wages as a measure to compare relative productivity differences within a country, the literature assumes that the proportion does not vary across regions within the country (Combes and Gobillon, 2015).

tributed across areas. They could also sort into such areas if they value cultural amenities. Second, the differences in productivity could be due to local non-human endowments. For instance, workers in some areas may have a higher marginal product because of geographical features such as a favorable location (like a port or a bridge on a river), a climate more suited to economic activity, or some natural resources. Arguably, local endowments cannot be restricted to natural features and should also encompass factors of production such as public or private capital, local institutions, and technology. Third, interactions between workers or firms take place locally and lead to productivity gains. Following Marshall (1890), denser input–output linkages between buyers and suppliers, better matching of workers’ skills with firms’ needs in thicker labour markets, and technological externalities resulting from more intense direct interactions are frequently mentioned as the sources of these gains (see Duranton and Puga (2004) for a review). A key issue is to understand whether these benefits stem from the size of the overall market (urbanisation economies) or from geographic concentration at the industry level (localisation economies). Understanding of the spatial differences in wages thus requires consideration of all these explanations. Addressing all of these factors simultaneously would allow understanding and quantifying the relevance of each explanation. Understanding the magnitudes associated with each explanation is especially crucial for formulation of policy to address these inequalities.

This paper aims to take up on the challenge of explaining the determinants of wages disparities across Turkish provinces by applying the “uniformed approach” proposed by Combes et al. (2008). To do so, I use social security records, a novel administrative dataset that covers the complete universe of private sector workers that are affiliated to social security, working in all of the industry. This is the first paper to use this new dataset.

In order to provide estimates that are comparable to the literature, I follow the standard two-step approach similar to Combes et al. (2008).¹⁸¹ This two-step approach allows me to distinguish gains due localisation economies (first-step) from those due to urbanization (second-step). Still, making causal interpretation of such forces requires addressing of two major identification issues: i) non-random spatial sorting of workers, ii) reverse causality from wages to city characteristics, if spatial wage differences are driving location choices of firms and workers.

The first concern requires accounting for spatial sorting of workers based on differences in their observable and unobservable characteristics. Addressing the sorting properly requires controlling of individual fixed effects through panel data. Such data however is unavailable in many developing countries, limiting the evidence from such contexts (Combes and Gobillon (2015)). I circumvent this problem by complementing my analysis by using the Household Labor Force survey, which provides individual-level data.

¹⁸¹As argued by Melo in their meta-analysis, differences in unit of analysis, the specification of agglomeration economies and the choice of controls can give rise to significant differences in results reported in the literature. Using a standard empirical approach is thus crucial to provide comparable estimates. The two-step approach was first proposed by Combes et al. (2008), later to be followed by Martin et al. (2011); Bakens et al. (2013); De la Roca and Puga (2017); Combes et al. (2019) among many others.

To address the endogeneity bias due to reverse causality, I use instruments based on past settlement patterns using historical data from the last census from Ottoman Empire and early census of the Turkish Republic.

I find the elasticity of wages with respect to employment density to be around 0.056-0.06.¹⁸² This means that doubling the employment density in an area increases the average wages by 3.8 - 4.2%. This elasticity is lower than one estimated for China (Combes et al., 2015; Chauvin et al., 2017) and India (Chauvin et al., 2017) and around those estimated for Brazil (Chauvin et al., 2017) and Colombia (Duranton, 2016). I also find a positive and strong effects for the domestic market potential. The estimated coefficient is around 0.091-0.1, which is double that of density, suggesting that having access to other markets is the most important determinant of the productivity differences in Turkey. This means that if the market potential of a province doubles (e.g., employment density doubles in all other provinces), the wages increase by 6.5%. This number is more than the triple of the 0.02 found for France in Combes et al. (2008), but smaller than 0.13-0.22 found for China in Combes et al. (2019).

I also do not find sorting of workers across locations according to their observable skills. This result is in sharp contrast with what is usually observed for developed countries, where a large fraction of the explanatory power of city effects arises from the sorting of workers (Combes and Gobillon, 2015). It is, however, very much in line with the results for China (Combes et al., 2015, 2019). This finding suggests that urbanisation patterns may be operating differently in developing countries, supporting the need for further evidence from such countries (Chauvin et al., 2017). Finally, I find a weak relationship between productivity (wages) and amenities, similar to the literature on developing countries. This pattern can be explained by either the high correlation between density and amenities (Duranton, 2016), or that workers in developing countries are not rich enough to forgo part of their income to live in areas with better amenities.

This paper contributes to the limited literature on the urban economics in developing countries by providing the first evidence from Turkey. It provides estimates from a highly urbanised, middle-income country and provide much needed evidence on the determinants of productivity in developing economies. I use a novel administrative data set, and use variety of panel data techniques and instrumental variables to deal with estimation concerns. I adopt a comprehensive and data intensive approach to provide estimates that are comparable with the rest of the literature. My findings corroborate earlier findings in the developing countries which show that while the main mechanisms of urban economies are present in the developing world, the current models need to be extended to capture the differences between the western cities and those in the developing part of the world.

The rest of the paper is organized as follows. In Section 5.2, I review the urban literature and the evidence from developing countries to prepare the ground for my empirical strategy and contextualise

¹⁸²Provinces (*il* in Turkish), correspond to the NUTS-3 (Nomenclature of Territorial Units for Statistics) level in the Eurostat classification of regions.

my findings. I then present the Turkish context (Section 5.3) and present the data used (5.5). In Section 5.6, I present my empirical strategy and discuss the identification issues. Section 5.7 provides estimates on density and other determinants of productivity. Section 5.10 concludes the paper.

5.2 Literature Review

To base my empirical approach, I start by providing an extensive literature review on the estimation of agglomeration economies (Section 5.2.1) before discussing the magnitudes found in the literature (Section 5.2.2) and in the developing countries (Section 5.2.3).

5.2.1 Agglomeration Economies

The idea that larger cities enjoy a productive advantage dates back to Marshall (1890), who argued that larger markets benefit from more intensive input-output linkages, thicker local labor markets, and technological spillovers between firms which in return increase the average productivity. To explain differences in productivity, the literature offers three broad explanations.

The first explanation attests that larger cities have higher productivity as their market size facilitates sharing, learning or matching (Duranton and Puga, 2004). In a seminal paper, Ciccone and Hall (1996) tested this hypothesis by measuring the size of the local economy through the number of individuals per unit of land, that is density. Using aggregate data for American states, they studied the impact of the logarithm of density on the logarithm of workers' productivity, measured by nominal wage. The use of density was crucial for two reasons: first, it allowed overcoming concerns about the mismeasurement of the size of the local economy due to the heterogeneity in the spatial extent of the geographic units that are used in these studies.¹⁸³ Second, it allowed capturing the correlation between higher concentration of activity and the productivity, that is the combination of both agglomeration economies and dispersion forces.¹⁸⁴

The second explanation puts forward that higher productivity is due to sorting of workers with higher abilities to larger cities. Glaeser and Maré (2001) was the first to test this idea by introducing individual fixed effects while studying the effect of density across US cities. The use of individual fixed effects was important as it allowed controlling for all the factors that can increase the individual's ability but remain constant over time. Since the individual fixed-effects can correlate flexibly with density and other local variables, it allows capturing the effects of local characteristics net of composition effects due to sorting on the individual characteristics.¹⁸⁵

¹⁸³The heterogeneity in the spatial extent of the geographic units and their shape (due to administrative borders) are problematic for two reasons. First, they make the comparison of estimates across studies difficult. Second, the size and the shape of the spatial units may influence the estimated elasticities due to "the modifiable areal unit problem". As shown by Briant et al. (2010), using density reduces concerns about both issues.

¹⁸⁴It should be noted that although density captures the net effect of spatial concentration, it does not allow identification of specific channel through which it operates (Combes and Gobillon (2015)).

¹⁸⁵This argument is further supported by Combes et al. (2008) who use an individual panel to estimate the effect of density on wages across all French cities. They show that not accounting for individual sorting (through fixed effects) leads

The agglomeration effects identified in the early literature focused on the instantaneous effect of density on productivity. These gains which are specific to the location and thus *static*, augment the productivity of workers as long as they are located in the area. However, larger cities can also generate *dynamic* gains as they provide faster learning opportunities (Lucas, 1988; Glaeser, 1999). Furthermore, these gains can have long lasting effects if workers are able to transfer part of their productivity gains from agglomeration across locations. This would mean that a worker who moves from a larger city to a smaller one can bring part of his productivity and be more productive than other individuals who have not worked in a large city.

Glaeser and Maré (2001) was the first to distinguish between the static and dynamic effects of agglomeration. Focusing on urban areas in Britain, D’Costa and Overman (2014) tested for both static and dynamic gains simultaneously. Using individual-level data for a large panel of British workers, they show that wage growth due to city size occurs in the first year that a worker moves to a city and this urban wage premium persists over time. In a more recent work, De la Roca and Puga (2017) proposed a dynamic framework that accounts for learning effects as a function of the time spent in different classes of city size. They argue that the knowledge that is acquired in larger cities is transferable over time and space. The degree of transferability of these dynamic gains, however, depends on the characteristics of locations.

5.2.2 Magnitudes for the Effect of Density on Workers’ Productivity

The contributions in the literature differ in terms of context, data, empirical models and identification strategies which make the comparison of their results and conclusions difficult (Melo et al., 2009). Regardless of these differences, it has now been established that the local density of economic activities increases the productivity of firms and workers. This conclusion emerges from a large number of studies across different periods, contexts, and data.¹⁸⁶

The earlier attempts to measure agglomeration economies used aggregate data, where the logarithm of regional wage (or TFP) would be regressed on the logarithm of employment or population density. Two benchmark studies using aggregate data for the United States are Ciccone and Hall (1996) and Rosenthal and Strange (2008), who found an elasticity of productivity with respect to density, at around 0.04–0.05. This elasticity range implies that the doubling of density increases the average productivity by 3 to 4%. Similar elasticities were estimated for European regions (Ciccone, 2002; Brühlhart and Mathys, 2008; Foster and Stehrer, 2009).

Glaeser and Maré (2001) were first to introduce individual fixed effects in estimation of the effect of

to the overestimation of the coefficients of the local variables.

¹⁸⁶Rosenthal and Strange (2004) and Combes and Gobillon (2015) provide an excellent overview of the literature. I only focus on recent articles that use richer datasets at the individual level that include workers’ or firms’ precise location, where the dependent variable is individual wages. See Combes and Gobillon (2015) for a review of the literature using TFP as a measure of productivity.

density on wages using data from the American cities. Using a much larger panel for workers, Combes et al. (2008) estimate the effect of density on wages across all French cities using individual fixed effects and also taking into account aggregate endogeneity using a two-step estimation procedure involving instrumentation. They find an elasticity of wages with respect to density of around 0.030, which is half the elasticity that is obtained when individual unobserved heterogeneity is not taken into account. Using approaches accounting for individual heterogeneities through fixed effects or controls, similar elasticities are found for European economies (i.e. 0.025 for Spain by De la Roca and Puga (2017)); 0.01 for Italy by Mion and Naticchioni (2009); 0.016 for Britain by D’Costa and Overman (2014); 0.021 for Netherlands by Groot et al. (2014)).

In order to assess the magnitude of dynamic effects, De la Roca and Puga (2017) consider a quantity defined at the city level as the sum of the time-invariant city fixed effect and the effect of experience accumulated in the city for a worker who stayed there for 7 years (which is the average length of time for workers in their sample). The elasticity of this measure with respect to density (which captures both static and dynamic effects) is 0.049, which is almost twice as large as the estimated elasticity of city fixed effect of 0.025 suggesting important dynamic gains.

The presence of reverse causality between productivity and agglomeration has been addressed through various instrumentation techniques. The most common strategy is to use long-lagged values of population or population density to instrument present values of agglomeration economies (Ciccone and Hall 1996; Rice et al. 2006; Combes et al. 2008; Mion and Naticchioni 2009; De la Roca and Puga 2017) and geographical instruments (Ciccone, 2002; Rosenthal and Strange, 2008; Combes et al., 2008). The motivation for the choice of these instruments is that both past levels of urban size and external geologic variables (e.g., soil composition, depth to rock, water capacity, soil erodibility, and seismic and landslide hazard) are correlated with current levels of urban size but not with current levels of productivity (Combes et al., 2010). Correcting for aggregate endogeneity has a small effect; sometimes decreasing the estimated elasticities by 10–20%, sometimes leaving the estimates unaffected or even increasing them slightly (Combes and Gobillon, 2015).

5.2.3 Evidence from Developing Countries

The literature from developing countries is recent and scant (Glaeser and Henderson, 2017). Furthermore, the lack of individual panel data in such countries makes it impossible to take into account unobserved individual heterogeneity. Differences between individuals have been taken into account through individual explanatory variables such as qualification, gender, age, and sometimes occupation or the type of firm where the individual is employed (Duranton, 2016; Chauvin et al., 2017; Combes et al., 2019).

Using individual data and standard instrumentation strategy, the effect of density on individual wages is found to be 0.05 in Colombia (Duranton (2016)), 0.09–0.12 in India (Chauvin et al., 2017) and 0.10–0.12

in China (Combes et al., 2015). Other measures of productivity have also been used in studies at the aggregate level, such as value added per worker in Korea (Henderson et al., 2001), establishment-level output per worker in Korea (Lee et al. (2010)), or output per worker in China (Au and Henderson, 2006), firm productivity in India (Lall et al., 2004) or in Chile (Saito and Gopinath, 2009).

The impact of market size on wages has been studied for China (Au and Henderson, 2006; Combes et al., 2013), India (Lall et al., 2004), and Colombia (Duranton, 2016).¹⁸⁷ Overall, these studies find that market size has a larger effect than in developed countries.

Finally, some articles have studied local determinants of agglomeration economies other than market size such as industrial specialisation (Henderson et al., 2001 in Korea; López and Südekum (2009) in Chile) and industrial diversity (Saito and Gopinath (2009) in Chile).

5.3 The Turkish Context

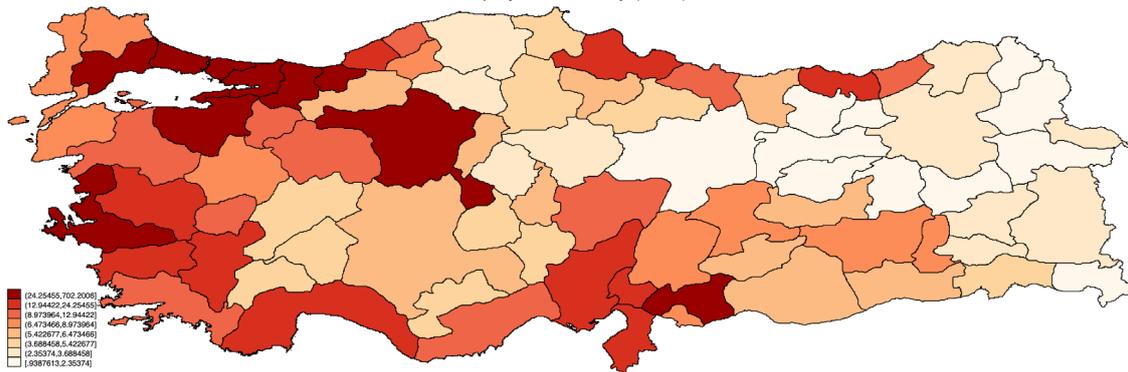
Turkey has a population of about 81 million over an area of 783 thousand square kilometers (Turkstat, 2019). While 74 percent of the country's population live in cities, some 92 percent of Turkey's gross value added is produced in cities (World Bank, 2015).

With a Gross Domestic Product (GDP) per capita of USD 10546 in 2017, Turkey is an upper-middle-income developing country (according to the World Bank classification). However, this wealth is not equally distributed across its regions (See Figure 37). While the GDP per capita in Istanbul was USD 17827, it was only USD 3489 in Ağrı (Turkstat, 2019). These differences are multiplied even further since population distribution is also uneven. In 2017, while 18.6% (15.1 million people) lived in Istanbul province, 6.7% (5.4 million) and 5.2% (4.3 million) lived in Ankara and Izmir, respectively. The population density of Istanbul, the densest province, is 2892 persons per square km, while it is only 11, in the least dense, Tunceli (Turkstat, 2019). All in all, while Istanbul is producing 31.5 percent of the national GDP, adding its immediate surrounding area increases the share to 39 percent in 2017 which creates large imbalances across regions in terms of production and income (Turkstat, 2017).¹⁸⁸

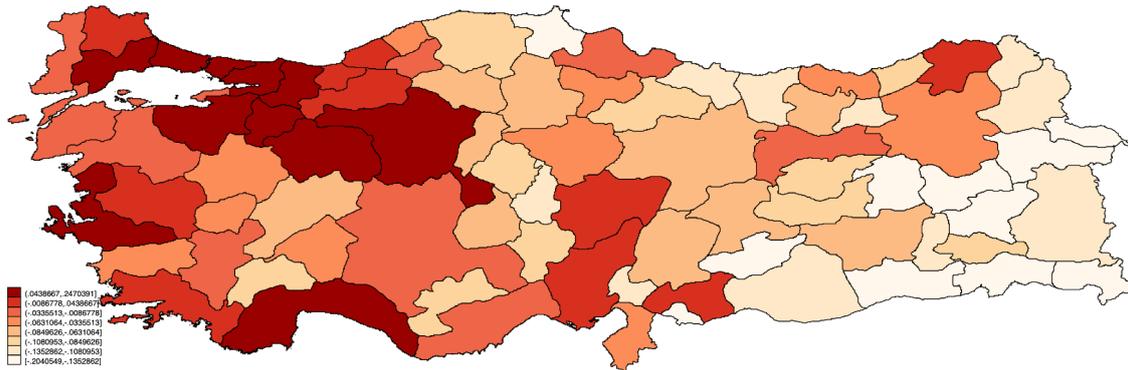
¹⁸⁷Similarly, the impact of market access on individual wages is also estimated in Brazil (Fally et al., 2010) and China (Hering and Poncet, 2010).

¹⁸⁸The surrounding includes Kırklareli, Tekirdağ, Kocaeli, Yalova, Sakarya and Bursa provinces.

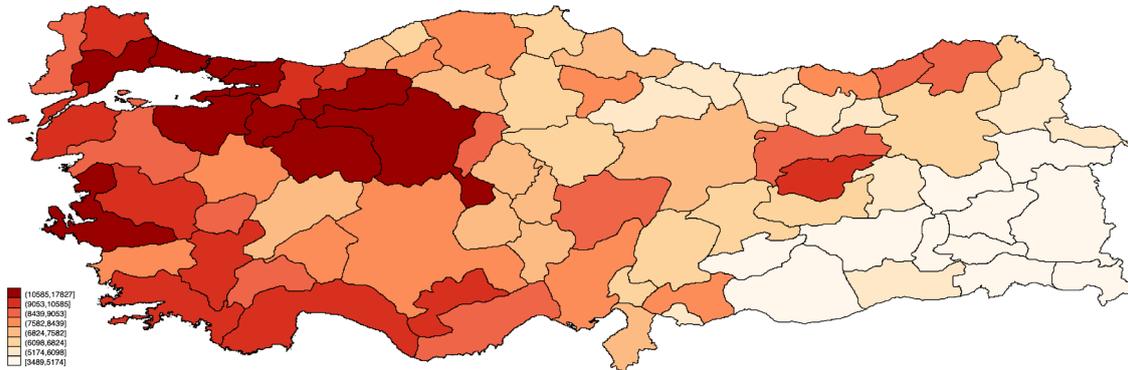
Figure 37: Spatial Differences
Employment Density (2013)



Productivity (2013)



GDP per capita in USD (2017)



Regional Imbalances Regional imbalances in Turkey go back to the late Ottoman Empire when the geographical location of Western Anatolia, especially the major ports of export like Izmir, Istanbul, and their hinterlands, gave these areas an essential role in the external trade of the country. Since then, trade and industry have always been more developed in these areas than in East Anatolia. With the foundation of the Republic of Turkey in 1923, the attention of the successive governments shifted to Central Anatolia where the capital, Ankara, was established.

In the early Republic Period (1923-1950) nation-wide industrialization and urbanization policies were developed to promote the planning and development of settlement areas, developing industries in major cities like Istanbul, Izmir, Ankara or Adana, but also in selected small Anatolian towns. Given

that the State was a majority owner of commercial activity, these plans generated significant spatial transformation in the economic activities and population.

Starting in 1963, addressing regional inequalities became an official priority of the State. The State Planning Organisation (*Devlet Planlama Teşkilatı*) was given the specific responsibility of reducing regional disparities through multi-annual planning that foresaw the public investments.¹⁸⁹ State manufacturing investments, public enterprises and transport investments aimed to expand development in the poorer parts of the country, especially in Eastern and Southeastern regions.

In the late 70s, the State's role as a technocratic agent of development shifted. During this period, the state began to recede from its interventionist mode to more of an enabler of the private sector (Boratav, 2017). In the third Development Plan (*Kalkınma Planı*, in Turkish) (1973-1977), Priority Provinces for Development (*Kalkınmada Öncelikli Yöreler*, in Turkish) all provinces of Eastern and Southeastern Anatolia were given priority in public investment. Going forward, all of the successive plans aimed to increase investment in these provinces both by increased public investment in the infrastructures and the offering of investment incentives to attract the private sector (such as tax break, lump-sum payments). In parallel, the State also ensured that all of the local administrative units received adequate financing capabilities.¹⁹⁰

After 1980, the new policies aimed at the development of the export base and favored the delocalization of the industrial activities from metropolitan cities to adjacent provinces of metropolitan regions. Regions located in the periphery of metropolitan cities started specializing in specific industries based on their comparative advantages in transportation network or natural resources. While the production moved to the periphery, Istanbul and other metropolitan areas, have increased its control on the private capital and management of the foreign trade (Gezici and Hewings, 2004). State continued also supporting private sector investment in less advanced regions and targeting public resources to carry out large infrastructure investments such as dams and new roads.

Urbanization In 1960, Turkey still featured a largely agrarian economy with 31 percent of its population residing in urban areas.¹⁹¹ While urbanization was steadily increasing during the 1950s-70s, it was during the 1980s that Turkey experienced a major surge of rural migrants to cities, causing rapid expansion of informal areas in urban settlements.¹⁹² During this period most of the rural migration were

¹⁸⁹According to Gezici and Hewings (2004), despite focusing on the regional inequalities, these plans were not effectively implemented and thus did generate a significant contribution to reducing inequalities. For a discussion on regional disparities in Turkey and government policies to tackle the issue see also Celebioglu and Dall'erna (2010).

¹⁹⁰To ensure that local administration units did not suffer from lack of financing, the Turkish state founded the Bank of Provinces (*İller Bankası*, in Turkish) in 1933. The Bank was mandated to provide long-term funding that was necessary for local administrative units to finance the prepared municipal development plans and infrastructure projects. To this today, this Bank remains the principal source of financing for municipalities.

¹⁹¹In terms of population, urban agglomerations were small in this period. Only one urban agglomeration (Istanbul) had more than 1 million, while 82 percent of the urban population lived in urban agglomerations with less than 500,000 people. Between 1965 and 1980, the number of urban agglomerations with 1 to 5 million people grew to 3. At that time, these large cities were home to about 39 percent of the country's population.

¹⁹²The mechanization of agriculture that began in the 1950s is one of the most significant transformations that took place in the modern Turkish economy (Pamuk, 2008) and triggered massive rural-urban migrations, which have profoundly

directed towards Turkey's three primary cities of Istanbul, Izmir, and the nation's capital of Ankara.

Between 2000 and 2010, Turkey's urban population has in a decade grown three times faster than its overall population. However, the growth during this period was driven by the country's secondary cities, which have experienced an increase of 53% of their population (World Bank, 2015). This growth was fueled by firms which were increasingly moving toward dynamic secondary cities, capturing economic spillovers from Turkey's large primary cities, while taking advantage of lower land rent values and labor costs. Turkey's principal cities, meanwhile, were diversifying their economies and focusing on innovation to remain competitive.

Overall since the 1960s, Turkey has structurally and demographically transformed from a predominantly agrarian economy to the globally competitive industrial economy it is today. During the country's most rapid period of urbanization, from 1960-2013, Turkey's industrial share of the economy increased from 17.6 percent to 27 percent, and the service sector dramatically rose from 26.4 to nearly 64 percent.

5.4 Literature on Regional Differences

Turkish regional imbalances has been the focus of a number of papers over the years. While some used provincial income data to study the spatial income inequalities (Atalik (1990); Gezici and Hewings (2004); Luca (2016)), others focused on provincial and regional convergence of income (Kirdar and Saracoglu, 2008; Celbis and de Crombrughe, 2018; Luca, 2016).

Although there has been a considerable number of studies dealing with the determinants of overall productivity (Altug et al., 2008; Ismihan and Metin-Ozcan, 2009; Atesagaoglu et al., 2017) or in the manufacturing sector (Krueger and Tuncer, 1982; Onder et al., 2003; Atiyas and Bakis, 2014), regional level analysis has been limited. Few examples include Metin et al. (2005) who study TFP change of the private and public sectors in the Turkish manufacturing industry in eighteen provinces from 1990-1998 or Temel et al. (1999) who use gross provincial product per worker for the period 1975-1990. These papers find evidence of concentration of high productivity activities in a few highly industrialized regions while most provinces tended to move towards low productivity activities, creating regional divergence in terms of productivity.

Determinants of spatial productivity in Turkey have also not been addressed in the literature. Coulibaly et al. (2007) is the closest work in spirit, to this paper. The authors assess the impact of urbanization on sectoral productivity between 1980 and 2000 by using manufacturing data and geographical, infrastructural and socio-economic data at province level. Their results suggest that localization (similar to specialisation which measures how much local production is concentrated in a given activity) and affected the Turkish demography. Most of the migrants were originating from provinces with a large share of the agricultural sector in the east moving to the more developed provinces in the west (Kirdar and Saracoglu, 2008). Many cities were unable to accommodate this growth, and the influx of migrants took place so quickly that these informal settlements became known as *gecekondu*, literally "houses erected overnight". States permissive policy toward rural migration and gecekondu encouraged migrants to flock to cities.

urbanization economies, as well as market accessibility increase productivity. Authors do not, however, deal with endogeneity of local determinants of agglomeration economies.

This paper is the first evidence from Turkey focusing on the spatial productivity differences through agglomeration literature perspective. The novel data that I use in this paper (see the next section), also makes it the first paper that analysis productivity differences at province level covering the period after the 2000s.

5.5 Data and Sample

This section presents different sources of data that are used to estimate the determinants of spatial differences in wages.

5.5.1 Social Security Data

In this paper, I use a novel administrative data set that was made available to researchers recently. This data set is collected by the Social Security Institution (*Türkiye Cumhuriyeti Sosyal Güvenlik Kurumu - SGK*) and are based on administrative records for all the workers affiliated to the social security system. It covers employment in all of the industries, in the private sector.¹⁹³

Due to data privacy issues, the raw individual-level data is aggregated by the SGK by sector, province and year. Thus the data includes yearly information on the number of workers, total number of days worked, number of firms and total payments received (wages and benefits) by the workers, for 81 provinces, grouped according to Nace Revision 2 at 4-digit sector level (659 sectors) for the period 2007-2013. The data is further disaggregated by job contract-type (temporary vs. permanent), by sex (male vs. female).

The novelty of this data is that it makes analysis at province-level possible. The literature focusing on individual or firm-level outcomes has been limited to analysis at region level (NUTS-2) due to data availability.¹⁹⁴ This dataset makes it possible, for the first time, to focus on productivity differences at a geographically more disaggregated level. This is also the first paper to use this dataset.

5.5.2 Household Labor Force Survey

I complement the main analysis by using individual level data obtained from the Household Labor Force Survey (*Hanehalkı İşgücü Anketi*, LFS henceforth) prepared by the Turkish Statistical Institute (*Türkiye*

¹⁹³The data covers all employment with compulsory insurance in the private sector under Article 4-1/a of Act 5510. For the year 2013, this corresponds to 12,5 million individuals. It does not include apprentices (321 thousand), those who work abroad but are affiliated with the Turkish Social Security System (35 thousand), the agricultural sector (64 thousand) and voluntary based insured partially employment (230 thousand). It does not include the self-employed (2.9 million) who are covered under Article 4-1/b of Act 5510, nor those who are employed in the public sector (2.8 million) under Article 4-1/c of Act 5510.

¹⁹⁴Two main data sources used in studies focusing on the Turkish labor market are Household Labor Force Survey (*Hanehalkı İşgücü Anketi*) and Annual Industry and Service Statistics (*Yıllık Sanayi ve Hizmet İstatistikleri*) which are provided by Turkish Statistical Institute (*Türkiye İstatistik Kurumu*). Both datasets allow identification at NUTS-2 level.

İstatistik Kurumu, Turkstat henceforth). The main objective of the labour force survey is to obtain information on the structure of the labour force in the country. The national labor force statistics, for instance, are produced based on the LFS surveys. It is representative of the total population in Turkey. It includes annual information on economic activity, occupation, status in employment and hours worked for employed persons (both formal and informal sector); unemployment, education and much more.

In the period of my analysis, each survey wave included around 135 thousand household, covering 500 thousand individuals. Its high level of detail and large sampling size makes it an important source of data in the labor market research in Turkey (e.g., Tumen, 2016; Balkan and Tumen, 2016; Baslevant and Onaran, 2003, 2004).

The main shortcoming of this dataset is that it allows geographical identification at regional level (NUTS-2). In my analysis, I prefer using the social security data which allows me to study the local interactions at a lower geographical unit (i.e., province-level, NUTS-3) and use the LFS as a complement.

5.5.3 Historical Data

Since Ciccone and Hall (1996), it is standard practice to use long-lagged variables as instruments for local characteristics. Following the literature, I construct various instruments using Ottoman Empire population statistics of 1914 and the Turkish Republic's population censuses of 1927, 1935 and 1945. The last Ottoman census was conducted in 1905/1906. Population statistics of 1914 is an updated version of this census. The 1914 population data used in this study were published for the first time by Karpat (1985), adapted to current administrative borders by Sakalli (2019).¹⁹⁵ I complement this data by digitizing published census reports for the period 1928, 1935 and 1945 which come from Turkstat. In addition to the population statistics, these data also include information on occupations, number of students, number of schools and much more.

I use these data to calculate the past population densities and past domestic market potential. I also use the number of enrolled male students to elementary schools and high schools in 1927. Finally I compute the foreign-market potential using historical GDP data coming from the Maddison Project Database (Bolt and van Zanden, 2014).

5.5.4 Controls

Land Area The data on the provincial land area come from Turkstat. It covers the real surface within the province borders, excluding lakes, in kilometer square.

Education Province level education data comes from Turkstat. The data is available at province level since 2008, includes the number of individuals by the highest diploma obtained (9 groups according to

¹⁹⁵Between 1914 and 2007, the number of districts, names, and their borders has changed considerably. I use the correspondence between past and current administrative boundaries prepared by Sakalli (2019).

International Standard Classification of Education (ISCED, henceforth) classification and by sex. I use this data to calculate the human capital measure which I explain in detail in Section 5.6.3.

Road Lengths I use real road lengths to measure distance between provinces. The data comes from the General Directorate of Highways (*Devlet Karayolları Müdürlüğü*, in Turkish). These distances measure the shortest route using real distances by road between the provincial centroids. I use this data to calculate market potentials. I also collected data on the length of provincial road and village road networks from Turkstat. These numbers reflect the sum of road network in kilometers (Km).

Foreign Market Potential I calculate the Foreign Market Potential using the GDP of the trading partners and bilateral distance between Turkey and other countries using the *Gravity database* provided by Centre d'Etudes Prospectives et d'Informations Internationales (CEPII) (Head et al., 2010).

Geographical Endowments Data on the distance from province centroid to the closest sea coast (*Coast*), length of seashores within the provincial boundaries (*Shore*), total length of rivers within the province (*River*) come from Sakalli (2019). See Sakalli (2019) for more details.

Cultural Amenities Data on amenities such as number of public libraries, cinema halls and public hospitals come from Turkstat. The data is provided annually for each province.

Weather All the data related to weather (i.e., mean average temperatures, average temperatures in january or june, average number of rainy days in a month and average number of sunny days in a month) is obtained from Turkish State Meteorological Service (*Devlet Meteoroloji Müdürlüğü*).

5.5.5 Sample

For my analysis I use data on workers from private sector, holding temporary and permanent job contracts, for both females and males. Thus my sample covers all the workers in the formal sector. I check the sensitivity of my results for subgroups as a robustness.

I focus on the period 2008-2013. I exclude 2007 as data on skill-levels is not available for that year. This leaves me with a final sample of 168 904 industry-province-year observations which I use to estimate the province-year fixed effects in the first step.

5.6 Empirical Strategy

This section presents the framework used for estimating the agglomeration effects in Turkey. It also discusses possible identification issues arising from i) missing variables, ii) reverse causality and iii) selection bias due to sorting of workers by their ability.

5.6.1 Econometric Equation

To evaluate the impact of agglomeration economies on the productivity, the following specification can be used:

$$\log w_{pst} = \alpha + \beta \log Den_{pt} + \gamma_s + \gamma_t + \varepsilon_{pst} \quad (38)$$

where w_{pst} is the average daily wage in province p , sector s at time t ,¹⁹⁶ den_{pt} is the total number of employees (or population¹⁹⁷) in province p at time t (emp_{pt}) divided by land area ($Area_p$).¹⁹⁸

Since Ciccone and Hall (1996), it is customary to measure the size of the local economy using density, which is the number of workers (or individuals) per unit of surface area. Although the number of employees can also be used directly, dividing over the land area is preferable as it addresses concerns due to heterogeneity in the spatial extent of the geographic units that are used. Moreover, the use of density addresses the concerns about the shape of the unit of analysis (due to the arbitrary administrative borders) which is known as the *modifiable areal unit problem* in the literature.¹⁹⁹ The β captures the total impact of local characteristics related to agglomeration economies rather than the magnitude of specific channels through which agglomeration forces operate. Moreover, it is the total of the net effects of density, which could be both positive and negative.

The panel structure of my data (81 provinces, 659 industries and 6 years), allows me to introduce sector fixed-effects (γ_s) which capture sector-specific differences in the productivity that is irrespective of time (e.g., average labor productivity is higher in manufacturing sectors than in agriculture), and time fixed-effects (γ_t) to absorb any temporal variations that affect the productivity of all provinces and sectors equally (e.g. productivity gains from technological progress). The use of time-fixed effects also addresses concerns due to the use of wages in current Turkish Lira and is more precise than using an arbitrary deflator (Combes and Gobillon, 2015). Moreover, nominal wages are a better of measure for capturing differences in the productivity compared to real wages, which would be capturing differences in “standard of living” (Duranton, 2016).²⁰⁰ Finally ε_{pst} is the error term representing unexplained productivity.

5.6.2 Estimation Issues

There are four sources of bias in the estimation of Equation 38. First, some other province characteristics may be correlated with density and wages. In that case employment density may capture the impact of

¹⁹⁶I calculate the daily wage by dividing the average monthly salary by the number of days worked.

¹⁹⁷Depending on the mechanism studied, agglomeration effects can be measured using employment, population, or production. Since these three variables are highly correlated separate identification of their impact is not possible. I use employment (instead of the population) as it reflects better the magnitude of the local economic activity (Combes and Gobillon (2015)). Moreover, given that some of the local variables (such as diversity or specialisation) can only be constructed using employment data, it is more consistent to measure the employment density.

¹⁹⁸Formally, this corresponds to: $Den_{pt} = emp_{pt}/area_p$.

¹⁹⁹For a discussion on why the use of density reduces the mismeasurement of the size of the local economy, see Briant et al. (2010)

²⁰⁰The estimation of local real wages requires considering the cost of living, specifically land prices. Given that such data is rarely available, nominal wages are used to have a consistent measure of productivity.

omitted variables that may be determining the productivity in the province. Second, workers may prefer to locate in cities where wages are higher. This creates a reverse causality problem. Third concern is related to the computation of standard errors. Fourth, workers may sort across locations based on their ability. In this section, I address these issues by proposing multivariate estimations to take multiple variables into account, and then to instrument some of them.

5.6.3 Estimation Issue 1: Omitted Variables

In order to address the endogeneity bias due to the omitted variables, I add a number of variables drawn from the economic geography literature.

Land Area The productivity can be impacted by the local density but also by the size of employment. If gains from agglomerations outweigh the costs that are associated, both the density and the size of the local economy can have a positive impact on the local productivity. In order to capture both effects, I add the province surface area, $area_p$. The impact of density, holding land area constant, reflects the gains from an increase in the number of workers in the province. It captures all the gains associated with a thicker labor market. The land area, holding density constant reflects the gains from increasing the spatial extent of the area (i.e., province). A larger area is likely to have more non-market interactions among agents than a smaller area as it is more populated. The data on the surface area of each province in Km^2 come from Turkstat.

Market Potential Agglomeration effects may be operating beyond the borders of the unit of analysis. For instance proximity to larger outlets can be a source of profitability, as better market access allows firms to export more and at a lower cost (Krugman, 1980). In other words, better market access implies a stronger demand for the output of local firms and this increases the value of the marginal product of their labor and can thus be expected to lead to an increase in local wages. Furthermore, better access to outside markets makes local firms more productive by allowing them to access a broader variety of goods at a cheaper price, which again should lead to higher wages (Krugman and Venables, 2006).²⁰¹

The most common approach in estimating the market potential is to construct a measure following Harris (1954):

$$MP_{pt} = \sum_{i \neq p} \frac{den_{pt}}{dist_{ip}}$$

where MP_{pt} is the market-potential of province p in time t , which is equal to the sum of densities in the surrounding areas, divided by the distance between each market (i) and the province (p). This measure assumes that trade and communications costs are proportional to the inverse of distance. Furthermore,

²⁰¹See Fally et al. (2010) for the importance of market access in wages in Brazil, and Hering and Poncet (2010) in China.

the own area is excluded (i.e., $i \neq p$) to avoid multicollinearity and also to identify separately the effects of the own location size (i.e. density) and external demand (i.e. market potential).

I calculate this variable by using employment density of the provinces (i.e. den_{pt})²⁰², divided by interprovincial bilateral distances provided by the General Directorate of Highways (*Devlet Karayolları Müdürlüğü*).²⁰³ These distances measure the shortest route using real distances by road between the provincial centroids.

Beyond the domestic demand coming from locations within the country, labor productivity may also be affected by differences in foreign market access. Being close to trading ports or borders can lower the trade costs for areas, allowing the firms to access international markets and benefit from larger demand and externalities that may improve their productivity. In order to capture the foreign market potential, I use a similar measure:

$$FMP_{pt} = \sum_c \frac{GDP_{ct}}{dist_{cp}}$$

where the foreign market-potential of a province, FMP_{pt} is measured as the sum of GDPs of all the trading partners²⁰⁴, divided by the distance between the capital of that country and the centroid of the province. The annual GDP data come from the World Bank. I constructed the distance measure by combining data from multiple sources. I used the bilateral distance between Turkey and other countries provided by CEPII to measure the international distance. Using Googlemaps, I calculated the real road distance between the province centroid and the closest international port or border crossing through which international trade of that province is likely to pass. Finally, I combined the domestic distance and the international distance to obtain the bilateral distance between a province and a trading partner.

Specialisation Industrial employment can generate productivity gains both when it is higher because total employment at the location is higher, and when the share of the industry is higher for given employment at the location. While the former would be contributing to the urbanization externalities, thus benefiting all the sectors, the latter would provide additional productivity gains that are specific to the industry. To decompose these effects, I measure the employment share of the industry within the local economy in the following way:

$$spe_{pst} = \frac{emp_{pst}}{emp_{pt}}$$

where the spe_{pst} measures the specialisation of the industry s , in province p at time t , as the share

²⁰²Instead of density, the sum of salaries can also be used to capture the size of the local market. My results are robust to the use of either of the measures.

²⁰³Using road distance is especially relevant in Turkey as the road transport is the primary mode of freight transportation in Turkey and accounts for 90% of domestic freight (by tonne-km) and passenger traffic (Cosar and Demir, 2016).

²⁰⁴An alternative to the GDP is to use the value of the total trade between Turkey and the partner countries. My results are robust to using either of the two measures, and can be provided if requested.

of employment of that industry (emp_{pst}) in the total employment in the location (emp_{pt}). This measure should be included along with the total employment (or employment density) for a precise interpretation (Combes, 2000). Both of these variables are expected to have a positive impact, when there are urbanization and localization economies.

Diversity The industrial diversity could also increase the local productivity by facilitating diffusion of knowledge within and between industries. This intuition made popular by Jane Jacobs (Jacobs, 1969) was formalized by Duranton and Puga (2001). Although various measures of diversity have been proposed, it is common to use the inverse of a Herfindahl index constructed from the shares of industries within local employment:

$$div_{pt} = \left[\left(\frac{emp_{pst}}{emp_{pt}} \right)^2 \right]^{-1}$$

where the div_{pt} is the degree of diversity in sectoral employment, in province p at time t , by the share of employment of that industry (emp_{pst}) in the total employment in the location (emp_{pt}).²⁰⁵ The measure captures the distribution of employment over all other industries and thus the urbanization economies that it generates.

Human capital externalities New Growth Theories emphasize the role of human capital as a determinant of productivity (Lucas, 1988). It is crucial to control for the skill-level in the location for two reasons. First, it is likely that locations with higher density also have greater share of skilled labor (Combes et al., 2008). If the differences in the skill levels are not controlled for, density will capture part of the gains due to compositional differences. Second, human capital accumulation generates benefits that go beyond the private gains to the worker and create positive externalities to those who are located in the area (Moretti, 1999). Introduction of control would allow for measuring the scale of these benefits.

The extent of human capital can be tested by adding a variable that captures the skills of the local labor force (Rauch, 1993; Moretti, 1999; Combes et al., 2011; Duranton, 2016). The linear approximation to a Cobb-Douglas production function requires adding the share of skilled labor as an explanatory variable:

$$HC_{pt} = \frac{Pop_{pt}^{hs}}{Pop_{pt}}$$

where HC_{pt} is a proxy for the share of skilled population in province p at year t , which is measured as the high-skilled population (Pop_{pt}^{hs}) with university degree or higher (ISCED4, ISCED5 and ISCED6)

²⁰⁵Combes and Gobillon (2015) recommend exclusion of own industry when calculating this measure for a clearer interpretation especially if it is used along with specialisation. However, when the number of industries is large, it makes little difference to drop the own industry from computation since the correlation between two measures obtained with and without the own industry is large. In my case, given the large number of industries, I do not drop the own industry.

over the total population (Pop_{pt}).

The coefficient is not an elasticity, so its interpretation should be different. Including this variable allows capturing the effect of human capital on productivity. More precisely, it captures the positive impact of an increase in the share of high-skilled workers for a given level of density (or population size). It is essential to underline that, given that I am using aggregate data, this variable captures both private and social gains due to human capital.

Amenities Local amenities determine both local density and wages. Productive amenities such as airports, transport infrastructures, and universities increase productivity and attract workers, which makes the density increase.²⁰⁶ These endowments can raise wages through various channels by lower exporting costs, cheaper supplies, or higher productivity. In that case, a positive bias in the estimated coefficient of density is also expected. Using a complete set of productive endowments would, however, raise serious endogeneity concerns (Combes and Gobillon, 2015).

As put forth by Roback (1982), consumption amenities such as cultural heritage, social life, or climate-related amenities can increase the attractiveness of some locations for workers and thus make density higher. Such amenities do not have any direct effect on productivity, but the increase in housing demand they induce makes land more expensive. As a result, local firms use less land relative to labor, and this decreases labor productivity when land and labor are imperfect substitutes. This causes a negative bias in the estimated coefficient of density since density is positively correlated with missing variables that decrease productivity. To avoid the bias, in my benchmark analysis I control for endowments that are exogenous such as distance to sea coast, length of rivers, and length of seashores and average temperatures.

In the final part of the paper, I extend further the number of controls and add climate-related amenities, which have the advantage of being exogenous to the local economy but also being well measured (Chauvin et al., 2017). Specifically, I include variables that account for average temperatures in January and its difference from the ideal temperature of 21.1 Celsius. I also add the average yearly number of sunny days and annual rain volume. All of the data is obtained from the Turkish State Meteorological Service (*Devlet Meteoroloji Müdürlüğü*).

Despite endogeneity concerns, I also provide additional results with other types of productive (e.g., road network) and non-productive amenities (e.g., cinema halls, hospitals, libraries) as it is informative. I construct all of these controls using data provided by Turkstat.

²⁰⁶ Arguably, local endowments cannot be restricted to natural features and should also encompass factors of production such as public or private capital, local institutions, and technology.

5.6.4 Estimation Issue 2: Circular Causality

The second estimation issue is that some local characteristics are likely to be endogenous to local wages. For instance, employment areas receiving a positive technology shock may attract migrants. This leads to a positive correlation between the residuals and the density of employment. In this particular case, reverse-causality is going to bias the estimates upwards. Other regressors such as market potential or human capital are likely to be endogenous since they also depend on workers' and firms' location decisions.

Several instrumentation strategies have been proposed in the literature to address this endogeneity issue (Combes and Gobillon, 2015). Using historical instruments, as proposed by Ciccone and Hall (1996), is the most popular method and it builds on the hypothesis that historical values of population (or density) are relevant for today's levels as they are persistent over very long periods. The local outcomes of today (such as productivity, types of economic activities), however, are unlikely to be related to the economic outcomes a long time ago that probably affected the historical population.

Following this strategy, I construct various instruments using Ottoman Empire population statistics of 1914 and the Turkish Republic's population census of 1927 and 1935. Using these historical population numbers, I build variables that capture population densities and population growth. The intuition is that current productivity shocks are not correlated with the employment structure from decades before the date of observation.

The instruments are valid in the case of Turkey for a few reasons. First, it is unlikely that density levels from almost 100 years ago to be correlated with labor productivity today, as the Turkish economy went through a wide range of productivity shocks during this period. Successive wars between 1914 and 1923 had a significant impact on physical and human capital stock while disrupting industrial and agricultural production in most parts of the country. In addition, considerable population shifts took place between 1914-1924, which caused a dramatic reduction in the share of employment in the non-agricultural sector. The urban population was disproportionately affected by the decade-long wars and their aftermath (Altug et al., 2008).

Economic government and policies have seen important changes as well. Following the transition from a multi-ethnic, multi-religious empire to a nation-state under a democratic and representation rule, the newly founded capital Ankara created policies that presented a contrast to the past. As discussed in Section 5.3, Turkey went through important sectoral re-allocation and experienced a significant structural transformation. Massive public investment in human capital increased literacy rates from around 10% in 1923 to 90% in 2007 (Altug et al., 2008).

I construct past domestic market potential using 1935 and 1945 population numbers, and past foreign market potential using the GDP levels in the main trading partners of Turkey in 1945 using the Maddison Project Database. In the main specification, I jointly use several instruments (instead of using only the

1914 urban population) for two reasons. First, since the population is taken in logarithmic form, using a multiplicity of census dates is equivalent to instrumenting by past levels and long-run historical growth rates. Second, having multiple instruments allows me to the instrument not only for employment density but also for the market potential, diversity, and even land area. I can also conduct exogeneity and over-identification tests. Although I use multiple instruments in the main estimations, I also provide robustness tests using various single instruments.

I instrument density, market potential, and human capital. Although endogeneity of local characteristics can be argued for almost all of them, I choose to instrument at most three of them simultaneously, as more than that would be extremely demanding in terms of identification power. I estimate these instrumented regressions either controlling for all or none of the non-instrumented variables and show that the results are consistent in both cases.²⁰⁷

Across the tables, I use two groups of historical variables to instrument for density. The first group consists of population density in 1914, and the population growth rate between 1914 and 1927. I have a broader set of possible instruments, and I experimented with many combinations, all of which yielded broadly consistent results with one another and with the OLS. I choose to be parsimonious and report estimations for different groups of workers using the same sets of instruments to allow for reliable comparisons.

5.6.5 Estimation Issue 3: Accounting for Local Shocks

Equation 38 can be estimated, including the controls, and the IV strategy explained before. However, this estimation would be problematic because it does not allow computing the variance of local shocks. This makes it impossible to distinguish local shocks from purely idiosyncratic shocks at the industry-location level, which is vital with missing endowment variables. Furthermore, in a single-step estimation, the variance of local shocks has to be ignored when computing the covariance matrix of estimators. This can create significant biases in the standard errors for the estimated coefficients of aggregate explanatory variables (Moulton, 1990). To address this problem Combes et al. (2008) propose a two-step estimation strategy which both solves this issue and has the advantage of corresponding to a more general framework. I estimate the following equations:

$$\log w_{pst} = \alpha + \delta \log Spe_{pst} + \gamma_s + \gamma_{pt} + \varepsilon_{pst} \quad (39)$$

$$\gamma_{pt} = \nu + \beta_1 \log Den_{pt} + \theta X_{pt} + \phi Z_p + \gamma_t + \varepsilon_{pt} \quad (40)$$

In the first step (Equation 39), I regress the log average daily wages ($\log w_{pst}$) in province p , sector

²⁰⁷For a similar argument see Combes et al. (2015, 2019).

s at time t on $\log Spe_{pst}$ which captures the effect of specialisation in a given sector on productivity, sector fixed-effect (γ_s) and province-year fixed-effect (γ_{pt}). The sector fixed-effects (γ_s) capture any sector-specific differences in the productivity that is irrespective of time, while location-year fixed effect can be interpreted as local wage indices after controlling for observed and unobserved industry effects. ε_{pst} is the error term.

The province-year fixed effects estimated in the first step are then used as the dependent variable in the second step (Equation 40) and regressed on local characteristics that impact the productivity levels. To account for the local structure I use density ($\log Den_{pt}$) but also X_{pt} , which includes time-varying controls (e.g. market potentials, diversity, human capital, road network, and more.) and Z_p , which includes the time-invariant controls (e.g., land area, temperatures, geographic controls, etc.). γ_t is the year-fixed effect which, as discussed before, takes care of any shock that impacts the productivity levels across the whole country, and also correct for the use of nominal wages which are not deflated.

This method is preferable for two reasons. First, as explained before, doing a two-stage estimation allows for estimating two separate error terms one for province-sector-year (ε_{pst}) and one for province-year (ε_{st}). This makes it possible, in a second step, to tackle the endogeneity of density and other location characteristics without addressing the sector-specific endogeneity issues, such as specialisation ($\log Spe_{pst}$).

Second, this procedure makes it possible to separately identify the localization economies (first-step) from those that are due to urbanization (second-step) as well. This is particularly important for policy formulation, as it helps determine whether policy focus should be on further developing existing sectors or encouraging the arrival of new activities to the region.

As a robustness check, in Appendix Section 5.11.3, I run a single-stage estimation and find very similar results.

5.6.6 Estimation Issue 4: Sorting by Ability

The final identification problem is the possible correlation between density and worker characteristics. If the workers' spatial distribution depends on their abilities, then the local productivity would also be affected by the differences in the composition of the workers. In the case of sorting based on ability, the estimated impact of local variables would be inflated as they would be capturing productivity gains also due to differences in the composition of the workers. These differences can be due to observed and unobserved ability.

The sorting of more-skilled workers into larger cities is observed in the US and Europe (Combes et al. 2008; Baum-Snow and Pavan 2012; De la Roca and Puga 2017) . In the case of the US, Baum-Snow and Pavan (2012) find sorting based on observables characteristics, yet none due to unobservable ones. In the case of France, however, Combes et al. (2008) show that controlling for observable skills is not enough to

remove the bias. Differently than these developed countries, Combes et al. (2015) find very weak sorting on observables in China. This weak relationship makes them conclude that in the absence of sorting based on observables, sorting due to unobservable characteristics is unlikely. These results show that both the degree of bias and its direct channel (i.e., observable or/and unobservable) are specific to each context.

The most commonly used strategy is to use a panel of workers and estimate the parameter by comparing the same workers across several locations as suggested by Combes et al. (2008). However, such data is hard to find for developing countries. An alternative solution is to control for an extensive set of individual characteristics to take care of differences in observable skills (Duranton 2016; Chauvin et al. 2017; Combes et al. 2019). If workers sort across locations based on their unobservable abilities, this method is not enough to fully address the potential bias.

If such bias exists in Turkey, then a positive correlation between average wages and city characteristics could reflect a composition effect due to the over-representation of more able workers in some provinces. Given that I am using data aggregated at industry-province level, the estimated coefficients would be inflated if there is sorting bias. It is thus important to see whether such bias exists in Turkey, and if so, measure its size. To measure the potential bias due to sorting, I use Household Labor Survey and exploit its individual-level dimension.

5.7 Main Results

In this section, I start by estimating the elasticity of wages to density by addressing gradually each identification concern discussed previously (Section 5.7.1). Then I present results of a multivariate framework accounting for local characteristics that determine the local productivity (Section 5.8). I end the section with estimations that include infrastructure and amenities (Section 5.9).

5.7.1 Density

I start with estimating the effect of density on average productivity. As discussed earlier, an unbiased estimate can be obtained if the coefficient is not suffering from identification issues explained previously. I explore the elasticity of density through a simple framework where these concerns are addressed separately, before moving to more complex models with multiple controls and instruments.

5.7.2 Omitted Variable Bias

I apply two-stage regression where in the first-stage I estimate province-fixed effects which are used as the dependent variable in the second stage.²⁰⁸ Table 61 presents OLS results for the second stage of

²⁰⁸Although the first-stage is not reported, all of the estimated province-fixed effects and specialisation are highly significant ($p < 0.001$). The estimated elasticity of log of specialisation is 0.065. This means that 1 percent increase in the specialisation of the local industry increases the average wages by 0.065 percent.

the estimation. Each column corresponds to a different model using observations over 6 years for 81 provinces. I use weights that are proportional to the number of observations used to compute the LHS variable as it corrects for heteroskedastic error terms and thereby achieve a more precise estimation of coefficients (Solon et al. (2015)). Standard errors are clustered at province-level.

Table 61: OLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Density	0.0624 (0.005)***	0.0700 (0.004)***	0.0506 (0.005)***	0.0608 (0.002)***	0.0449 (0.007)***	0.0504 (0.005)***	0.0614 (0.007)***
Land Area		0.0433 (0.018)**					0.0355 (0.010)***
Diversity			0.0767 (0.025)***				0.0170 (0.020)
DMP				0.0790 (0.014)***			0.0950 (0.023)***
FMP					0.2805 (0.086)***		-0.0592 (0.100)
Human Capital						5.7470 (1.338)***	3.0584 (0.869)***
N	486	486	486	486	486	486	486
Adj R2	0.83	0.86	0.85	0.88	0.85	0.87	0.94

The table reports OLS estimates for the impact of employment density in on average productivity. The unit of observations are provinces. Regressions are weighted by total employment in year. Standard errors are clustered at province-level. Data source: Turkstat, SGK
 $*p < 0.10, **p < 0.05, ***p < 0.01$

Column 1 of Table 61 shows that density has an elasticity of 0.06. This means that doubling the worker density increases the average productivity by 4%.²⁰⁹ If density of Iğdır (2.96, P25) were to increase to density of Mersin (11.3, P75), its productivity would increase by 14.7%.²¹⁰

In the following columns, I address the concern due to omitted variable bias by adding a number of controls that are standard in the literature.²¹¹ In Column 2, I add the land (surface) area of the province. The impact of land area is significant, which is in line with the literature. The positive coefficient suggests that, for a given density, a 1 percent increase in the land area, increases the average productivity by 0.04%. The elasticity of density ($\beta = 0.06$) increases slightly compared to Column 1, although given the standard errors, I cannot reject the null hypothesis that they are equal. In Column 3, I add diversity control, which is also very significant. The elasticity of density drops to 0.051, suggesting that provinces with higher employment density are also more diverse. In Column 4 and 5, I separately add controls to capture the domestic market potential (DMP) and the foreign market potential (FMP). Both controls

²⁰⁹ $2^{0.06} - 1 \approx 4\%$

²¹⁰This elasticity is the result of pooling observations across 6 years. However, as can be seen in Appendix Table 70, the coefficient remains stable when the regression is repeated separately by year.

²¹¹It is important to note by including these controls, I make the implicit assumption that the controls are exogenous. Furthermore, these regressions ignore the endogeneity concerns due to reverse causality and individual unobserved heterogeneity. I address both concerns in the following sections separately.

are highly significant, and they reduce the elasticity of density.

In Column 6, I account for human capital by adding the share of the population with a university degree or higher. The control is highly significant, and its inclusion reduces the coefficient of density which is expected given the high correlation between density and human capital levels. It is important to note, however, that human capital control is a share and is a semi-elasticity.

Finally, in Column 7, I include all of the controls to provide estimates that are comparable to the literature. When all controls are included, the coefficient of density is 0.0614 and is highly significant. This means that doubling the density of a location increases the average productivity by 4 percent, when all else is equal. While the land area remains significant, diversity turns insignificant which is in line with the literature (Combes and Gobillon, 2015). When regressed along with the domestic market potential, the foreign market potential becomes insignificant. This common problem is due to the high correlation between the two variables (≈ 0.93) and the lack of spatial variability of the foreign market demand given the way the variable is constructed.²¹² That is why, in the following sections, I work only with the domestic market potential. Lastly, the contribution of human capital remains powerful and significant.

5.7.3 Reverse Causality

The second important estimation issue is the reverse causality between density and wages. I address this concern and the potential bias it generates by implementing an IV strategy. This exercise is important for the estimation of an unbiased elasticity because a good instrument would take care of both the reverse causality and the omitted variable bias addressing both concerns simultaneously.

My main identification strategy consists in implementing the 2-stage-least-squares (2SLS) estimation outlined in Section 5.6. I use the estimated province-year fixed effects as dependent variable and include only density as an explanatory variable. In order to address the endogeneity of the main variables of interest, I use historical instruments that are based on historical census data from 1914, 1927, 1935 and 1945. Specifically, I instrument $\log Den_{pt}$ with population density in 1914 ($\log Den_{p1914}$), population density in 1927 ($\log Den_{p1927}$), employment density in 1935 ($\log Den_{p1935}$), employment density in 1945 ($\log Den_{p1945}$), and growth in population density between 1914 and 1927 ($DenGrowth_p$).

Figure 38 provides a visual representation of the first stage for density. It plots the density in 1914 in the horizontal axis against the average density in 2008-2013. Each observation in the figure corresponds to a province.

²¹²For more on the issue, see Redding and Venables (2004) and Combes et al. (2011).

Table 62: First Stage Regressions, 2008-2013

	(1)	(2)	(3)	(4)
	LnDen	LnDen	LnDen	LnDen
Density1914	1.4799 (0.197)***			
Density1927		1.1646 (0.110)***		
Density1935			1.7960 (0.171)***	
Density1945				0.0549 (0.003)***
N	468	486	342	378
Adj R2	0.76	0.80	0.81	0.85

The table reports OLS estimates for the impact of employment density in on average productivity. The unit of observations are provinces. The number of observations in Columns 3 and 4 is lower due to differences in the number of provinces. See footnote 35 for details. Regressions are weighted by total employment in year. Standard errors are clustered at province-level. Data source: SGK, Karpat(1985)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Each column presents the coefficients coming from regressions where I use a different instrument. In the first column of Table 62, I use the population density in 1914. The estimated coefficient is highly significant and around 1.17 which is similar to estimates reported in Combes et al. (2011). Specifically, an increase in the imputed past density by one percentage point leads to a 1.17 percentage point increase in the worker density between 2008-2013.

In the following columns, I repeat the exercise using densities from more recent years. All of the instruments regardless of the base year used or whether they capture employment or population density, pass the weak instrument test.²¹³ As expected the F-statistics gets larger when I use densities from more recent periods.

Table 63 presents the results for the estimation of Equation 40.²¹⁴ As explained previously, the parameter β_1 corresponds to the effect of employment density on the productivity levels of the provinces.

²¹³The number of observations in Columns 4 and 5 is lower due to differences in the number of provinces. Between 1923 and 2008, the number of provinces went from 57 to 81. The historical data in 1914 and 1927 were at district level, which allowed me to combine them according to the province boundaries in 2008. The data for 1935 and 1945, however, were only available at province-level. This made it impossible to distribute them according to the current number of provinces.

²¹⁴Appendix Table 68 reports the summary statistics for the second-stage estimation.

Table 63: 2SLS Results for 2008-2013

	OLS	Den 1914	Den 1927	Den 1935	Den 1945	Den Growth
Density	0.0624 (0.005)***	0.0560 (0.005)***	0.0582 (0.004)***	0.0547 (0.004)***	0.0537 (0.004)***	0.0632 (0.007)***
N	486	468	486	342	378	468
KP F-Stat		56.70	112.19	110.94	425.13	8.91

The table reports 2SLS estimates for the impact of employment density in on average productivity. The excluded instruments are reported at the header of each column. The unit of observations are provinces. The number of observations in Columns 3 and 4 is lower due to differences in the number of provinces. See footnote 35 for details. Regressions are weighted by total employment in year. Standard errors are clustered at province-level. Data source: SGK, Karpat(1985)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In the first column, I report the elasticity obtained through OLS estimation. In the following columns, I present 2SLS results where the variable density is instrumented with lagged densities that are reported in the header of each column. In the final column, I use the change in the population density between 1914 and 1927 as an additional instrument.

Few results stand out. First, regardless of the instrument, the elasticities remain stable and highly significant around 0.056-0.06. This suggests that the OLS estimates suffer from a positive bias around 10%, which is in line with the literature (Combes and Gobillon, 2015). Second, all of the instruments have strong first-stages, proving to be good predictors. It also shows that the results are not dependent on the use of a specific instrument and are thus robust. Third, the standard errors are very small, indicating high precision of the estimates. Finally, the elasticities are very similar to that found in the final column of Table 61 indicating that valid instruments can take care of both the bias due to reverse causality and the missing variables.

Table 61 shows that elasticity of productivity to density is between 0.056-0.06. This means that doubling the worker density increases the average productivity by 3.8 - 4%. This elasticity is comparable to those found in other countries that use similar specification. It is similar to 0.06 found in Combes et al. (2008) for French employment areas over the period 1976-1998, 0.05 in Ciccone (2002) for the five largest EU-15 countries at the end of the 1980s, 0.06 found in Ciccone and Hall (1996) for American counties in 1988.

Compared to estimates in other developing countries, the elasticity of density is slightly higher than 0.05 found in Colombia (Duranton, 2016), but lower than 0.09–0.12 found for India (Chauvin et al., 2017) or 0.10-0.12 found for China (Combes et al., 2013). Given the level of urbanization in Turkey, this elasticity fits precisely where it would be expected. However, these last papers use individual-level data, which allows them to net out the endogeneity due to the possible sorting of the higher ability individuals into denser areas. Thus their elasticities present gains purely due to agglomeration economies generated by the higher densities. As I am using aggregate data, the elasticity that I estimate however includes

the positive bias due to individuals sorting. This means that these elasticities can be biased up to 100% as shown, in the context of France, by Combes et al. (2008) although it is also possible for the bias not to exist if there is no correlation between individual characteristics (observed and unobserved) and local characteristics as found in China (Combes et al., 2015, 2017). I explore this issue further in the next section.

5.7.4 Sorting By Ability

As explained in Section 5.6.6, the parameters estimated in the second-step can be biased if workers with higher abilities sort into denser areas. Given that SGK data lacks the individual complement, it is impossible to test the existence of such bias and net out its effect if it exists.

As a way to detect the presence of such bias, I use Household Labor Force Survey. The survey is conducted every quarter to measure the state of the economy. It covers around 500 000 individual observations annually, including all ages and sex. It is sampled to be reflective of the Turkish population but also the state of the economy, as it is used for calculating unemployment rate and other labor market measures.

The main shortcoming of this data, given the objective of this paper, is that it is aggregated at NUTS-2 regional level. This means that the locations of individuals can be identified only at one of the 26 regions.²¹⁵ Use of smaller scale sizes are better for capturing benefits of interactions that decay with distance (Combes and Gobillon, 2015). Still, using larger local units may not be an important issue. According to Briant et al. (2010) using consistent empirical strategies (i.e. accounting for individual selection) largely reduces issues related to shape and size of the unit of analysis, and allow the estimation of unbiased estimates.

Different waves of the LFS are repeated cross-sections with no individual identifiers which makes it impossible to use individual fixed effects in my estimation similar to Combes et al. (2008). In order to examine the existence of selection, I follow Combes et al. (2015) and carry out two tests²¹⁶: first, I estimate the first step regressions by successively including different sets of explanatory variables, two, I compare the second step regression using fixed-year effects which were estimated in a first-step including in individual characteristics vs. those which were not.

First, I estimate the first step estimation including different sets of explanatory variables (location effects, individual characteristics, and firm characteristics) to understand their relative contribution to the log of the monthly wage. Table 64 reports the adjusted R^2 of each regression.²¹⁷

Individual characteristics (i.e., education, age, and sex) alone explain 41% of the variations in indi-

²¹⁵26 regions correspond to the NUTS-2 level. These regions have different geographical and population sizes. NUTS-2 regions such as Istanbul (TR10), Izmir (TR31) and Ankara (TR51) are identical to the NUTS-3 provincial borders. Other regions are formed by combining multiple provinces.

²¹⁶This approach is also used in Combes et al. (2019) for China, and Colombia in Duranton (2016).

²¹⁷Full estimation results are available if requested.

vidual wages. The explanatory power of firm characteristics (i.e., firm size) is 24%. Region dummies and specialisation together explain only 8%. These results suggest that individual characteristics are the main factors explaining individual wage disparities, followed by firm effects. The location effects and specialisation matter very little.

Table 64: Explanatory Power of Various Sets of Variables

Region effects	0.08
Individual characteristics	0.41
Firm characteristics	0.24
Region effects and individual characteristics	0.44
Region effects and firm characteristics	0.29
Individual and firm characteristics	0.45
All three sets	0.48
N	59078

Notes: Table presents Adjusted R-squares for individual wage regressions using data for 2008. Region effects include both region dummies and the specialisation variable. Data source: Household Survey (Turkstat), Karpat(1985)

These results reveal that these three sets of effects are fairly orthogonal. Region effects and individual characteristics together explain 44% of the wage disparities when the sum of their individual R^2 is 0.49. Similarly, while region and firm effects explain 29% of the variation, the sum of their individual R^2 is 0.32. Finally, individual and firm characteristics explain only 45% of the differences in wages, the sum of their individual R^2 is equal to 0.65. These results suggest that differences in observed wages cannot be attributed to differences in the composition of the labor force or the type of firms present. The absence of correlation between the effect of individual characteristics and region dummies suggest that workers do not sort across regions according to their observable characteristics. These results are in deep contrast with what is observed in developed countries where a significant fraction of the explanatory power of region effects arises from the sorting of workers (Combes and Gobillon, 2015). However, they are very similar to the findings of Combes et al. (2015) and Combes et al. (2019) for China.

To further examine the absence of sorting in Turkey, I carry out a second exercise. If individuals sort across regions according to their abilities, some local variables especially the density should be correlated with individual observables such as education or occupation. It is possible to test this by estimating the region-year fixed effects with and without individual characteristics in the first step.²¹⁸ If workers sort across locations by their observable characteristics, these estimated fixed effects should absorb them, and thus provide different results in the second step.

I use the waves for the Household Labor Force survey for the period 2008-2013, and create a sample that matches the SGK data. Similar to the SGK sample, I keep all male and female workers, who are

²¹⁸See Appendix Table 72 for the first-step results.

between 18-65 years of age, with positive income, employed in the private sector and affiliated to the social security system.²¹⁹ I drop self-employed as it could mean a large set of occupations (e.g., street vendors, shop owners) in a context like Turkey and are also not included in the SGK data.²²⁰ This leaves me with 412 137 individual observations over the 6 year period.

I use the two-step procedure with individual level data similar to Combes et al. (2008). The procedure consists in estimating the following specification:

$$\log w_{irst} = \alpha + \delta \log Spe_{rst} + \phi X_{it} + \gamma_s + \gamma_{rt} + \varepsilon_{irst} \quad (41)$$

$$\gamma_{rt} = \nu + \beta_1 \log Den_{rt} + \gamma_t + \varepsilon_{rt} \quad (42)$$

The first-step estimation of equation 41 evaluates the impact of individual i 's wage at year t , w_{irst} , of region-time fixed effects, γ_{rt} , for region r where worker i is employed at year t and region r 's specialisation in sector s where i is employed (for 88 Nace 2 industries), and a set of individual characteristics X_{it} , such as age, age squared, sex, education (7 groups), occupation (39 ISCO 88 categories). In the second-step, I use the estimated region-year fixed effect, γ_{rt} , as the dependent variable and regress it on region's employment density ($\log Den_{rt}$) and time-fixed effects, γ_t . I measure density using the survey data for consistency.²²¹

Table 65: Second Stage Regressions, 2008-2013

	No Controls		Individual Controls	
	OLS	2SLS	OLS	2SLS
Density	0.0616 (0.003)***	0.0560 (0.007)***	0.0534 (0.003)***	0.0512 (0.005)***
N	156	156	156	156
KP F-Stat		66.85		66.85

The table reports OLS and 2SLS estimates for the impact of employment density in on average productivity. The unit of observations are NUTS-2 Regions. Regressions are weighted by total employment in year. Standard errors are clustered at region-level. Data source: Household Survey (Turkstat), Karpas(1985)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

²¹⁹There are a few reasons I apply this condition. First, I drop workers in the informal sector to match it with SGK data in terms of coverage. Second most of the evidence in the literature use data in the formal sector (e.g., Combes et al., 2008; D'Costa and Overman, 2014; De la Roca and Puga, 2017). By focusing on formal employment allows me to provide numbers that are comparable with the literature. Still, in Appendix Section 5.11.5, I provide estimates including also workers in the informal sector.

²²⁰Although not reported, I tested the robustness of the results to make sure that they are not dependant on the sample selection. Inclusion of public sector employees slightly reduces the elasticity of density to 0.047-0.053, while the inclusion of those who are not affiliated to the social security increases the elasticity to 0.063-0.07. The last result is in line with Atesagaoglu et al. (2017) who argue that exclusion of informal sector causes an underestimation of productivity in the Turkish context. Inclusion of self-employed does not change elasticities. Results are available if requested.

²²¹In Appendix Table 73, I test the robustness of my estimate by measuring employment density using the SGK employment data. Results are almost identical.

Table 65 reports OLS and 2SLS results for the second-step.²²² In Columns 1 and 2, I regress the region-year fixed effects which are estimated in the first-step which includes only region-year dummies and specialisation control. In Columns 3 and 4, I regress region-year fixed effects, which are estimated in the first-step where I include also a set of individual characteristics. While the first stage is weighted with survey weights, the second stage is weighted with the number of workers used to estimate the region-year fixed effects in the first stage. All regressions include year fixed effects and errors are clustered at the region-level. In Columns 2 and 4, I instrument the current employment densities with the population density in 1914.

The elasticity of density remains stable across specifications, regardless of whether individual controls are included or not in the first step. The difference in the estimated elasticities in Columns 2 and 4, suggest that excluding individual controls only inflates the elasticity by 8%. This small number shows that the workers in Turkey do not sort across locations based on their observables.

As mentioned earlier, the sorting can also be based on the unobservable characteristics of the individual. However, as argued by Combes et al. (2015), it is very unlikely for a sorting based on unobservables to take place while sorting on observables is so weak. Granted, a final conclusion on the issue can only be given following an analysis using panel data. Such data however is currently unavailable in Turkey.

Second, the similarity between these findings and those found in China (Combes et al., 2015, 2019) could be indicative of some significant differences between developed and developing countries in terms of sources of productivity differences. While sorting based on individual abilities seems to be an important determinant of spatial wage differences in developed countries, this pattern does not seem to hold in Turkey or China. To explain the lack of sorting in China, Combes et al. (2015) argue that mobility restrictions due to the Hukou system could be preventing workers to sort across urban areas based on their ability. Such mobility restrictions do not exist in Turkey. On the other hand, Turkey has seen experienced massive rural-urban migration since the 1950s. These migrations waves were triggered by the mechanization of agriculture and ethnic conflict that has hit the southeast of Turkey since 1985. These massive migration moves were directed to bigger cities but mainly to the three big cities. For instance, between 1950 to 2008, Istanbul's population increased from 1.2 to 18 million. The arrival of such big waves of low-skilled workers originating from agricultural regions may have broken the link between urban externalities and the sorting.

²²²First stage results are presented in the Appendix Section 5.11.4. All the variables have the expected signs and are statistically significant. One result that is worth pointing is that the estimated elasticity of (log) specialisation is much smaller (0.0279) than the elasticities found in the main results using data aggregated at industry-location-year (0.06). The individual-level LFS data allow us to account for the individual characteristics (e.g. education, occupation) and estimate the effect of specialisation net of education and occupation. As sectoral choice and individual ability are highly correlated with industry characteristic, the estimated elasticity of specialisation using the aggregate data attributes part of the positive effect of ability on specialisation (and other variables that are measuring local characteristics). It is important to note, however, that part of the drop can also potentially be explained by the larger geographical scale. If localization benefits suffer from geographical decay, then it is reasonable to expect externalities due to specialisation to be weaker and the estimated coefficient to be smaller.

5.8 Multivariate Approach: A Unified Framework

As discussed in Section 5.1, spatial wage disparities can be explained in three broad categories (i.e., skills, endowments and interactions). Combes et al. (2008) propose a “unified framework” which includes all of these explanations to have a sense of the magnitudes of each contributing factor. Understanding the contribution of each factor is especially important to inform policy.

I estimate the Equation 40, including a set of controls that capture all of the explanations. Naturally, this exercise is demanding in terms of data and requires instrumenting of multiple variables simultaneously. As the exogeneity of economic geography variables is debatable, one needs to be cautious when including them. In my analysis, I introduce five local variables (Density, Domestic Market Potential, Human Capital, Land Area and Diversity). Similar to Combes et al. (2019) I instrument at most three of them simultaneously (*Density*, *Domestic Market Potential*, *Human Capital*), as more than that is demanding in terms of identification power. I start with estimating only with these instrumented variables and then include the other non-instrumented variables (*Land Area*, *Diversity*) and show that the results are consistent in both cases.

Specifically, I instrument $\log Den_{p,t}$ with population density in 1914 ($\log Den_{p1914}$) and growth in population density between 1914 and 1927 ($DenGrowth_p$); domestic market potential ($\log DMP_{pt}$) with domestic market potential in 1945 ($\log DMP_{p1945}$); and human capital (HC_{pt}) with number of enrolled male students in 1927 ($EnrolledMale_{p1927}$).²²³

In order to test the statistical relevance of these instruments, I report Cragg–Donald F-Statistic weak instrument test, Shea’s partial R^2 which shows that my instruments explain a large share of the variation in the instrumented variables, once potential inter-correlations among instruments have been accounted for, and finally, the Hansen J-Statistic tests over-identifying restrictions.

²²³As mentioned earlier, I have a larger set of possible instruments and I experimented with multiple combinations. Estimations using various combinations yielded largely consistent results with one another. I choose to be parsimonious, and report estimations using the same sets of instruments to allow for reliable comparisons. I also try to be restrictive about the number of instruments used and use just enough to carry out over-identification tests.

Table 66: 2SLS Results for 2008-2013

	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
Density	0.0680 (0.005)***	0.0641 (0.005)***	0.0642 (0.002)***	0.0609 (0.003)***	0.0518 (0.002)***	0.0541 (0.002)***	0.0629 (0.005)***	0.0720 (0.006)***	0.0580 (0.003)***	0.0635 (0.007)***
DMP			0.0904 (0.014)***	0.1002 (0.020)***	0.0813 (0.010)***	0.0871 (0.011)***	0.0816 (0.013)***	0.1059 (0.021)***	0.0956 (0.018)***	0.1059 (0.024)***
Human Capital					5.0698 (0.373)***	4.7589 (0.703)***	2.7917 (0.715)***	1.6219 (0.985)*	2.6255 (0.633)***	0.7702 (1.698)
Land Area							0.0286 (0.011)**	0.0403 (0.012)***	0.0312 (0.008)***	0.0389 (0.010)***
Diversity							-0.0032 (0.014)	-0.0267 (0.019)		
Shores									0.0019 (0.002)	0.0033 (0.003)
Coast									0.0399 (0.017)**	0.0504 (0.023)**
Climate									-0.0039 (0.001)***	-0.0061 (0.002)***
Rivers									0.0000 (0.000)	0.0000 (0.000)
N	486	468	486	468	486	468	486	468	486	468
r ²	0.87	0.86	0.92	0.91	0.95	0.95	0.96	0.95	0.97	0.96
Cragg-Donalds Stat		1161.06		443.73		131.58		68.02		33.83
P-value Hansen test		0.277		0.195		0.145		0.336		0.495
Shea's Partial(Density)		0.725		0.723		0.575		0.566		0.502
Shea's Partial(DMP)								0.512		0.582
Shea's Partial(HC)								0.619		0.673

The table reports OLS and 2SLS estimates for the impact of employment density and other controls on average productivity. The unit of observations are provinces. Regressions are weighted by total employment in year. Standard errors are clustered at province-level. Data source: Turkstat, SGK, Karpat(1985)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 66 presents the 2SLS results.²²⁴ For comparability, I start by presenting the OLS and 2SLS results where density is the only explanatory variable (Columns 1 and 2). In Column 3 and 4, I add domestic market potential, which I instrument with the domestic market potential in 1945. Introduction of this additional variable does not change the elasticity of density. In Columns 5 and 6, I include the human capital control which lowers the magnitude of both density and domestic market potential. This points to the relatively unequal distribution of the share of high skilled individuals, and a relatively strong correlation between employment density and human capital (Pearson's $R \approx 0.44$). In addition to the instruments used in the previous regressions, I add the number of enrolled male students in 1927 as an additional instrument. This positive and significant effect is expected as it captures both the private gains due to skills and the externalities generated by the presence of higher skilled individuals in an agglomeration.²²⁵

In Columns 7 and 8, I further control for the land area and diversity of the local economic activity. Although the former can be considered exogenous to the density, the latter is correlated. As the expected land area is highly significant, and for a given density, an increase in the land area increases the average productivity. If the land size of a province doubles, the wages increase around 3%. The diversity, on the other hand, is insignificant which is quite common in the literature (Combes and Gobillon, 2015).

Finally, I account for endowments. As discussed earlier, many productive endowments (such as airports, high-speed train lines, highways) can increase wages. However, given the endogeneity concerns in using such controls, I consider only four (exogenous) endowment variables that are related to the geography and thus are less concerning in terms of endogeneity. In Columns 9 and 10, I include controls to account for differences in length of seashores within the provincial boundaries (*Shores*), access to sea coast (*Coast*), mean annual temperature (*Climate*) and presence of rivers (*Rivers*).²²⁶ Compared to the previous columns, the inclusion of these controls does not impact the other coefficients. The coefficient of *Coast* is positive and significant, suggesting that having access to the shore improves productivity.²²⁷ While the *Shores* and *Rivers* do not seem to have any effect, the *Climate* seems to decrease productivity. Overall, the inclusion of these variables do not increase the explanatory power of the regressions (R^2) which is already high.

The last column (10) is my preferred specification, as it is the most comprehensive one. It includes controls that account for skills-based endowments (Human Capital), between-industry interactions (*Density*, *DMP*, *Human Capital*, *Land Area* and *Diversity*) and amenities (*Shore*, *Coast*, *Climate*, *Rivers*). *Density*, domestic market potential, and human capital are instrumented with long-lagged variables. The

²²⁴Appendix 75 table reports the first-stage results.

²²⁵The coefficient on the share of high skilled workers will also capture complementarities between skilled and unskilled labor in the production function. Also, as more educated workers flock to cities with higher wages, it also adds to the identification issues. Controlling for human capital make progress towards the identification of the true elasticity of density but also make it possible to compare it with the findings in the literature.

²²⁶I drop diversity as it is insignificant and generates unnecessary endogeneity.

²²⁷It is important to note that this variable captures the walking distance to the closest seashore. It does not imply having access to a port, which would be highly endogenous.

results pass all the relevant statistical test, and the model has high explanatory power.

The elasticity of density is 0.064, which is exactly the same as those found in Table 63. The domestic market potential remains positive and highly significant. The estimated coefficient (0.1) is a little less than the double of density, suggesting that having access to other markets is the most important determinant of the productivity differences. If the market potential of a province doubles (e.g., employment density doubles in all other provinces), the wages increase by 6.5%. This number is more than the triple of the 0.02 found for France in Combes et al. (2008), but smaller than 0.13-0.22 found for China in Combes et al. (2019).

5.9 Infrastructure and Other Amenities

In this section, I extend the number of controls used in the previous section and add three sets of variables that can impact the wages: transport infrastructure, cultural and additional set of climatic amenities. The average wages are also determined by the infrastructure and such amenities. The infrastructure amenities (e.g., road infrastructure, train-lines or airports) can improve the productivity by affecting the growth of urban areas, increasing trade and lowering cost of transportation (Redding and Turner, 2015). Although better infrastructures can improve overall productivity and increase wages, it can also reduce them through improving market access (thus lowering the prices of inputs and goods). The final effect on the wages is therefore, ambiguous. The literature on the effects of infrastructure is limited, mainly due endogeneity issue making it difficult to establish a causal effect.

Cultural amenities such as cinema halls, theaters or parks, can also impact the wages as they may increase the willingness of consumers to pay for land and thus imply higher local land rents (Roback, 1982). When the local prices increase, firms use relatively less land, which in turn can decrease the marginal product of labor, especially if the latter and land are not perfect substitutes. Similarly, amenities related to climate can also have a similar effect as it may increase the cost of land and living through higher housing prices and lower wages (Glaeser and Gottlieb, 2009).

Table 67: OLS Results for 2008-2013: Infrastructure and Amenities

	(1)	(2)	(3)	(4)	(5)
Density	0.0573 (0.004)***	0.0596 (0.003)***	0.0581 (0.006)***	0.0608 (0.003)***	0.0569 (0.006)***
Roads(KM)		0.0206 (0.028)			0.0116 (0.021)
Village Roads(KM)		-0.0142 (0.017)			0.0052 (0.019)
Cinema Halls			0.0085 (0.007)		0.0055 (0.007)
Hospitals			-0.0088 (0.016)		0.0107 (0.016)
Libraries			-0.0011 (0.018)		-0.0166 (0.013)
January				0.0011 (0.002)	0.0015 (0.002)
June				0.0048 (0.002)**	0.0055 (0.003)**
Sunny Days				0.0169 (0.005)***	0.0196 (0.007)**
Average Rain				-0.0001 (0.000)	-0.0000 (0.000)
Market Potential	Yes	Yes	Yes	Yes	Yes
Human Capital	Yes	Yes	Yes	Yes	Yes
Geographic Endowments	Yes	Yes	Yes	Yes	Yes
Land Area	Yes	Yes	Yes	Yes	Yes
N	486	486	486	486	486
Adj R2	0.95	0.95	0.95	0.95	0.96

The table reports OLS estimates for the impact of employment density in on average productivity. The excluded instruments are reported at the header of each column. The unit of observations are provinces. Regressions are weighted by total employment in year. Standard errors are clustered at province-level. Data source: SGK, Turkstat, WorldClim, Turkish State Meteorological Service

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 67 reports the OLS results where I augment the model in Section 5.8 with additional controls. In all of the regressions I control for market potential, human capital, geographic endowments, land area and include time-fixed effects. In Appendix Table 76, I also report the same results without these controls.

Column 1 replicates the results from Table 66 for comparison. In Column 2, I account for the length of provincial road network (*Roads*) and village roads (*Village Roads*). The results in the column suggest that denser road network does not impact the wages.²²⁸ As discussed earlier, better road networks can impact both positively and negatively the average wages. The non-significance of the results could be due to two opposing effects canceling each other out, and does not prove the absence of an effect.²²⁹It

²²⁸It should be noted that provincial and village roads are inferior to highways. While village roads are correlated with lower densities and rural economic structure, the relationship between provincial roads and density follows an inverse U-shape. That is why attempts to establish a linear relationships should be addressed with caution. Although not reported here, the highway network, which was very limited in the period of analysis, does not change results and remain insignificant.

²²⁹For more on the issue, see Duranton (2016).

should also be noted that these coefficients capture the effects that remain when controlling for domestic market potential.

In Column 3, I control for cultural amenities and health facilities. The number of cinema halls and libraries have the expected negative sign, although only the former is weakly significant (at 20% significance level). Hospitals, on the other hand, have a positive sign yet is insignificant. The interpretation of this sign is also should be done with caution. While better hospitals can increase the demand for the location, it can also be the consequence of higher levels of local income.²³⁰

In Column 4, I add controls to capture climate-related amenities following Chauvin et al. (2017). I account for average temperatures in January and June, and their difference from ideal temperatures of 21.11 Celcius.²³¹ I also control for the average number of sunny and rainy days in a month. The results suggest that Turks seem to get higher wages when they live in less temperate climates, although the magnitude is small. Despite significant differences in climate between Turkish provinces, it does not seem to matter for productivity or average wages.²³² This result is not surprising given that Turkey's economic divide goes from west-to-east, and other controls in the regression capture it. Finally, the provinces with a higher number of average sunny days seem to enjoy higher productivity levels. This result is possibly driven by the fact that Turkish provinces located on the Mediterranean coast have higher average incomes.

Given the endogeneity concerns and the difficulty in instrumenting, these OLS results should be viewed as providing robustness checks for my main findings. The absence of significant and robust effects are potentially due to the high correlation between density (and the other controls that are included) with the amenities. Although not presented, all of the amenities are correlated with density and other controls. The spatial inequalities are present in every measurable metric and thus are highly correlated with wages and average income in Turkey. This collinearity may prevent the appropriate identification of the effects of these amenities. For similar results found in the context of Colombia, Duranton (2016) argues that the collinearity is unlikely to be problematic when the standard errors for the amenities are small, which means that they are fairly precisely estimated. If true, this would indicate the weak relationship between productivity (wages) and amenities in a developing country context.

5.10 Concluding Remarks

This paper contributes to the literature on the agglomeration economies by providing evidence from Turkish provinces. Turkey is an excellent example of a developing country that has experienced fast

²³⁰Although not reported, I also tested for other amenities such as the number of theaters, museums, doctors, clinics, the number of beds in hospitals, and more.

²³¹The choice of 21.11 Celcius represents the middle ground between 18 and 24 degrees which is considered to be the ideal temperature for human comfort. Still, I check the robustness of this finding by using the similar measures obtained from Global Climate Data. Results can be provided if requested.

²³²This finding is in line with the literature (for Colombia (Duranton, 2016), or China and India (Chauvin et al., 2017) which shows that amenities (climate-related but also others) do not matter for developing countries.

urbanization and has a high share of the population living in urban areas. In addition to providing the first estimates on the determinants of spatial differences in productivity in Turkey, my findings also contribute to the broader knowledge base about agglomeration economies in developing countries.

Using a novel administrative dataset, employing various panel data techniques and instruments based on historical data, I find a positive and causal effect of density on productivity in Turkey. The estimated elasticity of 6 percent is higher than those estimated across U.S. and Europe, around those found for Colombia and Brazil, and smaller than those found for China and India. Consistent with the previous literature, I find a positive effect on market access, which is much stronger than those found in developed countries.

I also find evidence for very weak sorting of workers across provinces based on their observed abilities. This finding contrasts with what has been found for developed countries while echoing the findings for China. Put together, these findings hint that the models and stylized facts documented for cities in the developed countries may not apply fully to developing countries, thus requiring the extension of current models to match the realities in developing country contexts. Although this conclusion needs to be corroborated with further evidence coming from a broader range of developing countries, a better understanding of the sources of these differences are essential missing pieces in the literature, and thus remains high on my research agenda.

5.11 Appendix

5.11.1 Summary Statistics

Table 68: Summary Statistics for Local Variables

Variable	Mean	sd	p10	p25	p50	P75	P90
Density	18	66.4	1.8	3.0	5.7	11.3	24.9
LogDensity	2	1.1	0.6	1.1	1.7	2.4	3.2
Area (sq.km)	9,629	6,468.8	3,739.0	5,473.0	7,685.0	12,102.0	15,512.0
Diversity	37	14.3	21.1	27.2	34.9	44.5	55.3
Domestic Market Potential	3	1.6	1.3	1.7	2.4	3.3	4.9
Log Foreign Market Potential	23	0.1	23.1	23.2	23.3	23.4	23.5

Data source: Turkstat, SGK

Table 69: Summary Statistics for Individual Anaalysis

Variable	N	Mean	sd	Min	Max
Age	303,769	33.90	9.1	17.0	74.0
Female	303,769	0.23	0.4	0.0	1.0
Elementary School	303,769	0.02	0.1	0.0	1.0
High School	303,769	0.17	0.4	0.0	1.0
University	303,769	0.16	0.4	0.0	1.0

Data source: Turkstat

5.11.2 Yearly OLS results

Table 70: Yearly Regressions

	2008	2009	2010	2011	2012	2013
Density	0.0651 (0.004)***	0.0657 (0.004)***	0.0628 (0.005)***	0.0614 (0.005)***	0.0611 (0.004)***	0.0588 (0.006)***
N	81	81	81	81	81	81
Adj R2	0.85	0.85	0.82	0.82	0.84	0.78

The table reports OLS estimates for the impact of employment density in on average productivity. The unit of observations are provinces. Regressions are weighted by total employment in year. Standard errors are clustered at province-level. Data source: Turkstat, SGK

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.11.3 One-step results

Table 71 reports the results from one-step regression. I regress the log of the average wage in industry s , in province p at time t , on local characteristics and specialisation. The elasticities are almost identical to the main results obtained from two-step regression. The very high F-stats and low R^2 are due to the low variability since the dependent variable varies at the industry-province-year level while the independent variables vary only at the province-year level.

Table 71: One Step Regressions, 2008-2013

	OLS	2SLS	OLS	2SLS
Density	0.0713 (0.004)***	0.0685 (0.007)***	0.0584 (0.003)***	0.0667 (0.010)***
DMP			0.0600 (0.016)***	0.0682 (0.017)***
HC			6.7508 (0.519)***	4.1024 (1.932)**
Land Area			0.0399 (0.008)***	0.0559 (0.018)***
Diversity			0.0121 (0.007)*	0.0110 (0.018)
Specialisation			0.0759 (0.006)***	0.0762 (0.006)***
N	175333	171755	175333	171755
r2	0.64	0.18	0.69	0.30
Cragg-Donalds Stat		513986.18		18881.31

The table reports 2SLS for the one-step of the estimation. The dependent variable is the log of average earnings in industry s , in province p at year t . The unit of observations are industry-province pairs. Excluded instruments are the same as in the second-step of the estimation in the main text. All regressions include industry and year fixed effects. Regressions are weighted by number of workers in each industry-province cell. Standard errors are clustered at region-level.

Data source: SGK, Karpat(1985), Turkstat

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.11.4 Individual Analysis

First step individual regression

Table 72: First Step Regressions, 2008-2013

	Log(Earnings)
Age	0.0607 (0.000)***
Age2	-0.0007 (0.000)***
Female	-0.1360 (0.001)***
Elementary School	-0.0129 (0.004)***
Middle School	0.0719 (0.005)***
High School	0.1593 (0.005)***
Vocational School	0.1796 (0.005)***
University	0.3921 (0.005)***
Specialisation	0.0279 (0.001)***
Occupation FE (39 groups)	Yes
Industry FE (87 groups)	Yes
RegionXYear FE (156 groups)	Yes
N	303769
Adj R2	0.60

The table reports OLS for the first step of the estimation. The dependent variable is log of earnings, regressed on individual characteristics, region-year fixed effects (not reported) and industry-fixed effects (not reported). The unit of observations are individuals. Regressions are weighted by survey weights. Standard errors are clustered at region-level. Data source: Household Survey (Turkstat)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Robustness of the second-step

In Section 5.7.4, I exploit the individual dimension of the LFS to study the potential bias due to the sorting of workers across locations based on their observable and unobservable characteristics. In the second step of the estimation, I regress the region-year fixed effects estimated in the first step, on the worker density of the region, which I compute also using the LFS for consistency. As a robustness, I repeat this estimation using density that is calculated using the SGK data. Table 73 presents these results.

Table 73: Second Stage Regressions, 2008-2013

	No Controls		Individual Controls	
	OLS	2SLS	OLS	2SLS
Density	0.0580 (0.003)***	0.0527 (0.007)***	0.0507 (0.002)***	0.0483 (0.004)***
N	156	156	156	156
KP F-Stat		76.17		76.17

The table reports OLS and 2SLS estimates for the impact of employment density in on average productivity. The unit of observations are NUTS-2 Regions. Regressions are weighted by number of workers used in the first stage. Standard errors are clustered at region-level. Data source: Household Survey (Turkstat), SGK, Karpat(1985)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.11.5 Informal Employment

In developing economies, informal labor accounts for a substantial share of both urban and rural employment. According to Turkstat, the share of informal employment in the Turkish labor market stands high at 38.4 percent as of January 2012 (TurkStat, 2012). Moreover, TurkStat reports that the rate of informality to be 82.8 for the agricultural employment and 25.8 percent for the nonagricultural employment.

In this paper, I focus on the agglomeration economies based on formal employment. Given the prevalence of the informal sector, it is important to discuss the potential biases that could arise due to the omission of informal employment.

Given the objective of this paper, the exclusion of informal sector may matter for two reasons. First, it is important to understand the correlation between the share of informal employment and the density. If the distribution of informal employment as a share of the total employment is not homogenous across areas of different population sizes, this could create a bias in the estimated elasticity. Second, it can create omitted variable bias. I address these concerns in this section.

I use the Household Labor Force Survey as it allows distinguishing between workers who are affiliated to the social security vs. those who are not. Using this information I calculate the share of informal employment in the total employment in the region. The correlation between employment density (in levels) and share of informal employment is -0.22 (p-value<0.001). This suggests that areas with higher employment density have lower share of informal employment.

Furthermore, I estimate the impact of density on average wages including the workers employed in the informal sector. Specifically, I estimate the area-year fixed effects (in the first step) and the density (in the second step) including workers who are not affiliated to the social security. Table 74 presents second-step results. In Columns 1 and 2, I present the OLS and 2SLS results using only workers employed in the formal sector, for comparison.²³³ In Columns 3 and 4, I also include the workers who are in the informal sector.

²³³I estimate the first step controlling for a full set of individual characteristics. These results correspond to the results in Column 4 in Table 65.

Table 74: Accounting for informal sector, 2008-2013

	OLS	2SLS	OLS	2SLS
Density (Formal)	0.0499 (0.003)***	0.0481 (0.005)***		
Density (Formal+Informal)			0.0638 (0.004)***	0.0631 (0.006)***
N	156	156	156	156
KP F-Stat		76.20		90.86

The table reports OLS and 2SLS estimates for the impact of employment density in on average productivity. The unit of observations are NUTS-2 Regions. Regressions are weighted by number of workers used in the first stage. Standard errors are clustered at region-level. Results in Columns 1 and 2 are obtained only using workers in the formal sector, while results in Columns 3 and 4 include also include workers with informal employment. Data source: Household Survey (Turkstat), SGK, Karpat(1985)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Results show that the inclusion of informal sector increases the estimated coefficient to 0.064. This suggests that i) the agglomeration economies matter for the informal sector, ii) these effects are stronger for the informal sector. In this sense, the elasticities estimated (using formal employment) in the paper present the lower bound estimates of the agglomeration economies. Although understanding the source of the difference in the elasticities is interesting, it goes beyond the scope of this paper.

5.11.6 First Stage of Multivariate Regression

Table 75 reports first stage results for the Table 66. Column titles refer to the endogenous variables that are instrumented. Titles above the results refer to the column numbers in Table 66.

Table 75: First Stage Regressions, 2008-2013

	Column 2		Column 4		Column 6		Column 8		Column 10		Column 12	
	LnDen	LnDen	LnDMP	LnDen	LnDMP	ISCED56	LnDen	LnDMP	ISCED56	LnDen	LnDMP	ISCED56
Density1914	1.2162 (0.188)***	1.2123 (0.187)***	0.0218 (0.019)	0.6511 (0.393)	0.0181 (0.069)	-0.0041 (0.003)	0.2469 (0.414)	-0.1479 (0.087)*	-0.0008 (0.002)	0.3689 (0.340)	-0.1193 (0.080)	-0.0010 (0.001)
DenGrowth	0.0031 (0.001)***	0.0032 (0.001)***	-0.0004 (0.000)**	0.0014 (0.001)	-0.0004 (0.000)**	0.0000 (0.000)	0.0006 (0.001)	-0.0007 (0.000)***	0.0000 (0.000)*	0.0015 (0.001)	-0.0004 (0.000)*	0.0000 (0.000)
DMP1945		-0.5908 (0.689)	1.7268 (0.146)***	0.2631 (0.738)	1.7325 (0.143)***	0.0063 (0.004)	-0.1578 (0.668)	1.6712 (0.124)***	0.0045 (0.004)	0.1803 (1.138)	1.6422 (0.216)***	0.0087 (0.005)
(sum) enrolled _{m,ale} 1927				0.0000 (0.000)*	0.0000 (0.000)	0.0000 (0.000)**	0.0000 (0.000)**	0.0000 (0.000)	0.0000 (0.000)*	0.0000 (0.000)	0.0000 (0.000)	0.0000 (0.000)*
Land Area							-0.4412 (0.349)	-0.1727 (0.050)***	0.0032 (0.002)*	-0.1200 (0.294)	-0.0892 (0.049)*	0.0024 (0.001)*
Diversity							0.9593 (0.456)**	0.2239 (0.083)***	0.0000 (0.001)	0.8798 (0.407)**	0.2310 (0.075)***	0.0001 (0.001)
Shores										0.0230 (0.056)	-0.0015 (0.010)	0.0006 (0.000)*
Coast										0.9531 (0.488)*	-0.0629 (0.115)	0.0052 (0.002)**
Climate										0.0361 (0.073)	0.0195 (0.015)	-0.0002 (0.000)
Rivers										-0.0007 (0.001)	-0.0007 (0.000)***	0.0000 (0.000)
N	468	468	468	468	468	468	468	468	468	468	468	468
Adj R2	0.81	0.81	0.76	0.83	0.76	0.58	0.85	0.80	0.63	0.87	0.83	0.68

The table reports first-stage results. Each column reports regression results where the variable on the column head is the dependent variable, regressed on various variables and controls. All the regressions include year-fixed effects and standard errors are clustered at province-level. Data source: Turkstat, Karpat (1985)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.11.7 Additional Results: Infrastructure and Amenities

Table 76: OLS Results for 2008-2013: Infrastructure and Amenities

	(1)	(2)	(3)	(4)	(5)
Density	0.0625 (0.004)***	0.0758 (0.003)***	0.0564 (0.013)***	0.0654 (0.005)***	0.0946 (0.014)***
Roads(KM)		0.0866 (0.022)***			0.1142 (0.019)***
Village Roads(KM)		-0.0529 (0.017)***			-0.0386 (0.026)
Cinema Halls			0.0363 (0.012)***		0.0170 (0.009)*
Hospitals			-0.0422 (0.026)		-0.0441 (0.026)*
Libraries			0.0197 (0.019)		-0.0173 (0.016)
January				0.0011 (0.002)	0.0026 (0.002)*
June				-0.0111 (0.006)*	-0.0006 (0.004)
Sunny Days				0.0171 (0.010)	-0.0096 (0.009)
Average Rain				-0.0004 (0.000)	0.0002 (0.000)
N	486	486	486	486	486
Adj R2	0.82	0.88	0.85	0.84	0.90

The table reports OLS estimates for the impact of employment density in on average productivity. The excluded instruments are reported at the header of each column. The unit of observations are provinces. Regressions are weighted by total employment in year. Standard errors are clustered at province-level. Data source: SGK, Turkstat, Turkish State Meteorological Service

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

References

- Adserà, Alícia, and Mariola Pytliková. “The Role of Language in Shaping International Migration.” *Economic Journal* 125, 586: (2015) F49–F81.
- Ager, Philipp, and Markus Brückner. “Cultural diversity and economic growth: Evidence from the US during the age of mass migration.” *European Economic Review* 64: (2013) 76–97.
- Akin, Ozlem, José García Montalvo, Jaume García Villar, José Luis Peydró, and Josep Maria Raya. “The real estate and credit bubble: evidence from Spain.” *SERIEs* 5, 2-3: (2014) 223–243.
- Alesina, Alberto, Johann Harnoss, and Hillel Rapoport. *Birthplace diversity and economic prosperity*, volume 21. Springer US, 2016.
- Altonji, Joseph G., and David Card. “The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives.” *NBER Working Paper* Immigratio, January: (1991) 201–234.
- Altug, Sumru, Alpay Filiztekin, and Sevket Pamuk. “Sources of Long-Term Economic Growth for Turkey.” *European Review of Economic History* 12, 3: (2008) 393–430.
- Amior, Michael. “The Contribution of Foreign Migrants to Local Labor Market Adjustment.” 972, November: (2017) 1–75.
- Amuedo-Dorantes, Catalina, Cristina Borra CPD, San Diego State University, and Cristina Borra. “Internal Mobility after the Expansion of the Welfare State: Evidence from Spain.” .
- Amuedo-Dorantes, Catalina, and Sara De La Rica. “Immigrants’ Responsiveness to Labor Market Conditions and Its Implications on Regional Disparities: Evidence from Spain.” .
- Amuedo-Dorantes, Catalina, and Sara de la Rica. “Complements or substitutes? Task specialization by gender and nativity in Spain.” *Labour Economics* 18, 5: (2011) 697–707.
- Angrist, Joshua D., and Adriana D. Kugler. “Protective or counter-productive? Labour market institutions and the effect of immigration on EU natives.” *Economic Journal* 113, 488: (2003) F302–F331.
- Antolin, Pablo, and Olympia Bover. “Regional Migration in Spain: The Effect of Personal Characteristics and of Unemployment, Wage and House Price Differentials Using Pooled Cross-Sections.” *Oxford Bulletin of Economics and Statistics* 59, 2: (1997) 215–235.
- Antràs, Pol, Davin Chor, Thibault Fally, and Russell Hillberry. “Measuring the upstreamness of production and trade flows.” *American Economic Review* 102, 3: (2012) 412–416.
- Arango, Joaquín. “Exceptional in Europe? Spain’s Experience with immigration and Integration.” Technical report, 2013.

- . “Spain new emigration policies needed and emerging diaspora.” *Migration Policy Institute* 24.
- Arrow, Kenneth J. “Classificatory Notes on the Production and Transmission of Technological Knowledge.” *The American Economic Review* 59, 2: (1969) 29–35.
- Atalik, Gündüz. “Some effects of regional differentiation on integration in the European community.” *Papers of the Regional Science Association* 69, 1: (1990) 11–19.
- Atesagaoglu, Orhan Erem, Ceyhun Elgin, and Oguz Oztunali. “TFP growth in Turkey revisited: The effect of informal sector.” *Central Bank Review* 17, 1: (2017) 11–17.
- Atiyas, I, and O Bakis. “Aggregate and sectoral TFP growth in Turkey: a growth accounting exercise.” *Iktisat, Isletme ve Finans* 29, 341: (2014) 9–36.
- Au, Chun Chung, and J. Vernon Henderson. “Are Chinese cities too small?” *Review of Economic Studies* 73, 3: (2006) 549–576.
- Autor, David H, David Dorn, and Gordon H Hanson. “The Geography of Trade and Technology Shocks in the United States.” In *American Economic Review*. 2013, volume 103, 220–225.
- Autor, David H., David Dorn, and Gordon H. Hanson. “The China Shock: Learning from Labor-Market Adjustment to Large Changes in Trade.” *Annual Review of Economics* 8, 1: (2016) 205–240.
- Autor, David H., and Mark G. Duggan. “The rise in the disability rolls and the decline in unemployment.”, 2003.
- Autor, David H, Frank Levy, and Richard J Murnane. “The Skill Content of Recent Technological Change: An Empirical Exploration.” *The Quarterly Journal of Economics* 118, 4: (2003) 1279–1333.
- Bahar, Dany, Ricardo Hausmann, and Cesar A. Hidalgo. “Neighbors and the evolution of the comparative advantage of nations: Evidence of international knowledge diffusion?” *Journal of International Economics* 92, 1: (2014) 111–123.
- Bahar, Dany, and Hillel Rapoport. “Migration, Knowledge Diffusion and the Comparative Advantage of Nations.” *Economic Journal* 128, 612: (2018) F273–F305.
- Bakens, Jessie, Peter Mulder, and Peter Nijkamp. “Economic impacts of cultural diversity in the netherlands: Productivity, utility, and sorting.” *Journal of Regional Science* 53, 1: (2013) 8–36.
- Balkan, Binnur, and Semih Tumen. “Immigration and prices: quasi-experimental evidence from Syrian refugees in Turkey.” *Journal of Population Economics* 29, 3: (2016) 657–686.
- Bartik, Alexander W. “Worker Adjustment to Changes in Labor Demand : Evidence from Longitudinal Census Data.” *Mimeo* .

- Bartik, Timothy J. *Who Benefits from State and Local Economic Development Policies?* 1991.
- Baslevent, Cem, and Ozlem Onaran. “Are married women in Turkey more likely to become added or discouraged workers?” *Labour* 17, 3: (2003) 439–458.
- . “The Effect of Export-Oriented Growth on Female Labor Market Outcomes in Turkey.” *World Development* 32, 8: (2004) 1375–1393.
- Basso, Gaetano, D’Amuri Francesco, and Giovanni Peri. “Immigrants, Labor Market Dynamics and Adjustment to Shocks in the Euro Area.” *NBER Working Paper* .
- Baum-Snow, Nathaniel, and Ronni Pavan. “Understanding the city size wage gap.” *Review of Economic Studies* 79, 1: (2012) 88–127.
- Becker, Randy A, Wayne B Gray, and Jordan Marvakov. “NBER-CES Manufacturing Industry Database: Technical Notes.” Technical Report February, 2013.
- Becker, Sascha O., and Andreas Ferrara. “Consequences of forced migration: A survey of recent findings.” *Labour Economics* , February.
- Bell, Martin, Elin Charles-Edwards, Philipp Ueffing, John Stillwell, Marek Kupiszewski, and Dorota Kupiszewska. “Internal Migration and Development: Comparing Migration Intensities Around the World.” *Population and Development Review* 41, 1: (2015) 33–58.
- Bentolila, Samuel, Pierre Cahuc, Juan J. Dolado, and Thomas Le Barbanchon. “Two-Tier Labor Markets in the Great Recession: France vs. Spain.” *Economic Journal* 122, August: (2012a) 155–187.
- Bentolila, Samuel, and Juan J Dolado. “Mismatch and Internal Migration in Spain , 1962-1986.” In *Mismatch and Labour Mobility*, edited by Padoa-Schioppa F, Cambridge: Cambridge University Press, 1991, 1962–1986.
- Bentolila, Samuel, Juan J. Dolado, Wolfgang Franz, and Christopher Pissarides. “Labour Flexibility and Wages: Lessons from Spain.” *Economic Policy* 9, 18: (1994) 53.
- Bentolila, Samuel, Juan J. Dolado, and Juan F. Jimeno. “Reforming an insider-outsider labor market: the Spanish experience.” *IZA Journal of European Labor Studies* 1, 1: (2012b) 1–29.
- Bentolila, Samuel, and Andrea Ichino. “Unemployment and consumption near and far away from the Mediterranean.” *Journal of Population Economics* 21, 2: (2008) 255–280.
- Bentolila, Samuel, Marcel Jansen, and Gabriel Jiménez. “When Credit Dries Up: Job Losses in the Great Recession.” Technical report, 2017.

- Bertoli, S, J Fernández-Huertas Moraga, and F Ortega. “Crossing the Border: Self-selection, Earnings and Individual Migration Decisions.” *Journal of Development Economics* 101, 1: (2013) 75–91.
- Bertoli, Simone, and Jesús Fernández-Huertas Moraga. “Multilateral resistance to migration.” *Journal of Development Economics* 102: (2013) 79–100.
- . “The size of the cliff at the border.” *Regional Science and Urban Economics* 51: (2015) 1–6.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. “How Much Should We Trust Difference-in-Difference Estimates?” *The Quarterly Journal of . . .* 119 (1), February: (2004) 249–275.
- Besley, T., and R. Burgess. “Can Labor Regulation Hinder Economic Performance? Evidence from India.” *The Quarterly Journal of Economics* 119, 1: (2004) 91–134.
- Black, Richard. “Return and Reconstruction in Bosnia-Herzegovina: Missing Link, or Mistaken Priority?” *SAIS Review* 21, 2: (2001) 177–199.
- Black, Richard, and Saskia Gent. “Sustainable Return in Post-conflict Contexts . retours durables au lendemain d’un conflit. retorno sostenible en contextos posconflictos.” *International Migration* 44, 3: (2006) 15–38.
- Blanchard, Olivier Jean, and Lawrence F. Katz. “Regional Evolutions.” *Brookings Papers on Economic Activity* 1992, 1: (1992) 1.
- Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen. “What Drives Differences in Management Practices?” *American Economic Review* 109, 5: (2019) 1648–1683.
- Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. “Does management matter? Evidence from India.” *The Quarterly Journal of Economics* 128, 1: (2013) 1–51.
- Bloom, Nicholas, Kalina Manova, John van Reenen, Stephen Teng Sun, and Zhihong Yu. “Managing trade: Evidence from China and the US.” *NBER Working Paper Series* .
- Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. “Americans do IT better: US multinationals and the productivity miracle.” *American Economic Review* 102, 1: (2012) 167–201.
- Bloom, Nicholas, and John Van Reenen. “Measuring and Explaining Management Practices Across Firms and Countries.” *Quarterly Journal of Economics* CXXII, November: (2007) 1351–1408.
- Blossfeld, Hans-peter. “Labor-Market Entry and the Sexual Segregation of Careers in the Federal Republic of Germany.” *American Journal of Sociology* 93, 1: (1987) 89–118.

- Bolt, Jutta, and Jan Luiten van Zanden. “The Maddison Project: Collaborative research on historical national accounts.” *Economic History Review* 67, 3: (2014) 627–651.
- Bonhomme, Stéphane, and Laura Hospido. “The Cycle of Earnings Inequality: Evidence from Spanish Social Security Data.” *The Economic Journal* 127, 603: (2017) 1244–1278.
- Boratav, Korkut. *Türkiye İktisat Tarihi 1908-2015*. Istanbul: Imge Kitabevi Yayinlari, 2017.
- Borjas, G. J. “The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market.” *The Quarterly Journal of Economics* 118, 4: (2003) 1335–1374.
- Borjas, George J. “Does Immigration Grease the Wheels of the Labor Market?” *Brookings Papers on Economic Activity* , 1: (2001) 69–133.
- Borjas, George J. “Native internal migration and the labor market impact of immigration.” *Journal of Human Resources* 41, 2: (2006) 221–258.
- Borjas, George J., and Abdurrahman Aydemir. “Cross-Country Variation in the Impact of International Migration: Canada, Mexico, and the United States.” *Journal of the European Economic Association* 5, 4: (2007) 663–708.
- Borjas, George J, Jeffrey Grogger, and Gordon H Hanson. “Comment: On Estimating Elasticities of Substitution.” *Journal of the European Economic Association* 10, 1: (2012) 198–210.
- Bosswick, Wolfgang. “Development of Asylum Policy in Germany.” *Journal of Refugee Studies* 13, 1: (2000) 43–60.
- Bound, John, and Harry J. Holzer. “Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s.” *Journal of Labor Economics* 18, 1: (2002) 20–54.
- Boustan, Leah Platt, Price V. V. Fishback, and Shawn Kantor. “The Effect of Internal Migration on Local Labor Markets: American Cities during the Great Depression.” *Journal of Labor Economics* 28, 4: (2010) 719–746.
- Bover, Olympia, and Pilar Velilla. “Migration in Spain: Historical Background and Current Trends.” In *European Migration: What do we know?*, edited by Klaus F. Zimmermann, Oxford: CEPR and Oxford University Press, 2005, 389–414.
- Briant, Anthony, Pierre-Philippe Combes, and Mirene Lafourcade. “Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations?” *Journal of Urban Economics* 67, 3: (2010) 287–302.
- Brühlhart, Marius, and Nicole A. Mathys. “Sectoral agglomeration economies in a panel of European regions.” *Regional Science and Urban Economics* 38, 4: (2008) 348–362.

- Cadena, Brian C. “Native Competition and Low-Skilled Immigrant Inflows.”, 2010.
- Cadena, Brian C, and Brian K Kovak. “Immigrants Equilibrate Local Labor Markets: Evidence from the Great Recession.” *American Economic Journal: Applied Economics* 8, 1: (2016) 257–290.
- Cadot, Olivier, Céline Carrère, and Vanessa Strauss-Kahn. “Export Diversification: What’s behind the Hump?” *Review of Economics and Statistics* 93, 2: (2011) 590–605.
- Card, David. “Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration.” *Journal of Labor Economics* 19, 1: (2001) 22–64.
- . “Immigration and Inequality.” 1–21.
- Card, David, and John DiNardo. “Do Immigrant Inflows Lead to Native Outflows?” *American Economic Review* 90, 2: (2000) 360–367.
- Card, David, Jörg Heining, and Patrick Kline. “Workplace Heterogeneity and the Rise of West German Wage Inequality*.” *The Quarterly Journal of Economics* 128, 3: (2013) 967–1015.
- Card, David, and Francis Kramarz. “Changes in the relative structure of wages and employment: A comparison of the United States, Canada, and France.” *Canadian Journal Of Economics* 32, 4: (1996) 843–877.
- Card, David, and Giovanni Peri. “Immigration Economics by George J. Borjas: A Review Essay.” *Journal of Economic Literature* 54, 4.
- Carlino, Gerald, and William R. Kerr. *Agglomeration and Innovation*, volume 5. Elsevier B.V., 2015, 1 edition.
- Carrasco, Raquel, Juan F. Jimeno, and A. Carolina Ortega. “The Effect of Immigration on the Labor Market Performance of Native-born Workers: Some Evidence for Spain.” *Journal of Population Economics* 21, 3: (2008) 627–648.
- Celbis, Mehmet Güney, and Denis de Crombrughe. “Internet infrastructure and regional convergence: Evidence from Turkey.” *Papers in Regional Science* 97, 2: (2018) 387–409.
- Celebioglu, Fatih, and Sandy Dall’erba. “Spatial disparities across the regions of Turkey: An exploratory spatial data analysis.” *Annals of Regional Science* 45, 2: (2010) 379–400.
- Cette, Gilbert, Simon Corde, and Rémy Lecat. “Firm-level productivity dispersion and convergence.” *Economics Letters* 166: (2018) 76–78.
- Chauvin, Juan Pablo, Edward Glaeser, Yueran Ma, and Kristina Tobio. “What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States.” *Journal of Urban Economics* 98: (2017) 17–49.

- Choudhury, Prithwiraj. “Return migration and geography of innovation in MNEs: A natural experiment of knowledge production by local workers reporting to return migrants.” *Journal of Economic Geography* 16, 3: (2016) 585–610.
- Ciccone, Antonio. “Agglomeration effects in South-South FDI.” *European Economic Review* 46: (2002) 213–227.
- Ciccone, By Antonio, and Robert E Hall. “Productivity and the Density of Economic Activity.” *The American Economic Review* 86, 1: (1996) 54–70.
- Clemens, Michael, and Jennifer Hunt. “The Labor Market Effects of Refugee Waves: Reconciling Conflicting Results.” Technical report, 2017.
- Clemens, Michael A. “Economics and emigration: Trillion-dollar bills on the sidewalk?” *Journal of Economic Perspectives* 25, 3: (2011) 83–106.
- Clemens, Michael A, Ethan G Lewis, and Hannah M Postel. “Immigration restrictions as active labor market policy: Evidence from the Mexican Bracero exclusion.”, 2018.
- Combes, Pierre-Philippe. “Economic Structure and Local Growth: France, 1984-1993.” *Journal of Urban Economics* 47, 3: (2000) 329–355.
- Combes, Pierre-Philippe, Sylvie Démurger, and Shi Li. “Urbanisation and migration externalities in China.” .
- . “Migration externalities in Chinese cities.” *European Economic Review* 76: (2015) 152–167.
- . “Productivity gains from agglomeration and migration in the people’s Republic of China between 2002 and 2013.” *Asian Development Review* 34, 2: (2017) 184–200.
- Combes, Pierre-Philippe, Sylvie Démurger, Shi Li, and Jianguo Wang. “Unequal migration and urbanisation gains in China.” *Journal of Development Economics* .
- Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. “Spatial wage disparities: Sorting matters!” *Journal of Urban Economics* 63, 2: (2008) 723–742.
- Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux. “Estimating agglomeration economies with history, geology, and worker effects.” In *[Book] - Agglomeration Economics*, edited by Edward Glaeser, Cambridge, MA: NBER, 2010, volume I, 15–65.
- Combes, Pierre-Philippe, and Laurent Gobillon. *The Empirics of Agglomeration Economies*, volume 5. Elsevier B.V., 2015, 1 edition.

- Combes, Pierre-Philippe, Miren Lafourcade, JF Jacques Francois Thisse, and Jean Claude Toutain. “The rise and fall of spatial inequalities in France: A long-run perspective.” *Explorations in Economic History* 48, 2: (2011) 243–271.
- Cortés, Patricia, and Jessica Pan. “When Time Binds: Substitutes for Household Production, Returns to Working Long Hours, and the Skilled Gender Wage Gap.” *Journal of Labor Economics* 37, 2: (2018) 351–398.
- Cosar, Kerem, and Banu Demir. “Domestic road infrastructure and international trade: Evidence from Turkey.” *Journal of Development Economics* 118: (2016) 232–244.
- Costinot, Arnaud, Dave Donaldson, and Ivana Komunjer. “What Goods Do Countries Trade? A Quantitative Exploration of Ricardo’s Ideas.” *Review of Economic Studies* 79: (2012) 581–608.
- Coulibaly, Souleymane, Uwe Deichmann, and Somik Lall. “Urbanization and Productivity: Evidence from Turkish Provinces over the Period 1980-2000.” *World Bank*, August: (2007) 32.
- Cuñat, Alejandro, and Marc J Melitz. “Volatility, labor market flexibility, and the pattern of comparative advantage.” *Journal of the European Economic Association* 10, 2: (2012) 225–254.
- Dao, Mai, Davide Furceri, and Prakash Loungani. “Regional labor market adjustment in the United States: Trend and cycle.” *Review of Economics and Statistics* 99, 2: (2017) 243–257.
- Dasgupta, Kunal. “Learning and knowledge diffusion in a global economy.” *Journal of International Economics* 87, 2: (2012) 323–336.
- Dauth, Wolfgang, Sebastian Findeisen, and Jens Suedekum. “The Rise Of The East And The Far East: German Labor Markets And Trade Integration.” *Journal of the European Economic Association* 12, 6: (2014) 1643–1675.
- David, S Reher, and Silvestre Javier. “Internal Migration Patterns of Foreign-Born Immigrants in a Country of Recent Mass Immigration: Evidence from New Micro Data for Spain1.” *International Migration Review* 43, 4: (2009) 815–849.
- D’Costa, Sabine, and Henry G. Overman. “The urban wage growth premium: Sorting or learning?” *Regional Science and Urban Economics* 48: (2014) 168–179.
- De Chaisemartin, C., and X. D’Haultfoeuille. “Fuzzy differences-in-differences.” *Review of Economic Studies* 85, 2: (2018) 999–1028.
- De La Rica, Sara, Albrecht Glitz, and Francesc Ortega. “Immigration in Europe: Trends, policies, and empirical evidence.” In *Handbook of the Economics of International Migration*. 2015.

- De la Roca, Jorge. “Wage cyclicality: Evidence from Spain using social security data.” *SERIEs* 5, 2-3: (2014) 173–195.
- . “Selection in initial and return migration: Evidence from moves across Spanish cities.” *Journal of Urban Economics* 100: (2017) 33–53.
- De la Roca, Jorge, and Diego Puga. “Learning by working in big cities.” *Review of Economic Studies* 84, 1: (2017) 106–142.
- Dengler, Katharina, Britta Matthes, and Wiebke Paulus. “Occupational Tasks in the German Labour Market : an alternative measurement on the basis of an expert database.” Fdz methodenreport, Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg, 2014a.
- . “Occupational Tasks in the German Labour Market : an alternative measurement on the basis of an expert database.” Fdz methodenreport, Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg, 2014b.
- Deutscher-Bundestag. “Bericht der Beauftragten der Bundesregierung für die Belange der Ausländer über die Lage der Ausländer in der Bundesrepublik Deutschland 1993. BT-Drs. 12/6960.” <http://dipbt.bundestag.de/doc/btd/12/069/1206960.pdf> (22.09.2017), 1994.
- Dimova, Rozita. “From Protection to Ordeal: Duldung Status and Bosnians in Germany.” *Max Planck Institute for Social Anthropology Working Papers* 49, 87: (2006) 0–12.
- Dix-Carneiro, Rafael, and Brian K Kovak. “Trade liberalization and regional dynamics.” *American Economic Review* 107, 10: (2017) 2908–2946.
- Docquier, Frédéric, Caglar Ozden, and Giovanni Peri. “The labour market effects of immigration and emigration in OECD countries.” *Economic Journal* 124, 579: (2014) 1106–1145.
- Duranton, Gilles. “Agglomeration effects in Colombia.” *Journal of Regional Science* 56, 2: (2016) 210–238.
- Duranton, Gilles, and Diego Puga. “Nursery cities: Urban diversity, process innovation, and the life cycle of products.” *American Economic Review* 91, 5: (2001) 1454–1477.
- . “Micro-Foundations of urban agglomeration economies.” In *Handbook of Regional and Urban Economics*, edited by J.Vernon Henderson, and JF Jacques Francois Thisse, Elsevier, 2004.
- Dustmann, Christian, Tommaso Frattini, and Ian P. Preston. “The effect of immigration along the distribution of wages.” *Review of Economic Studies* 80, 1: (2013) 145–173.
- Dustmann, Christian, Tommaso Frattini, and Anna Rosso. “The effect of emigration from poland on polish wages.” *Scandinavian Journal of Economics* 117, 2: (2015) 522–564.

- Dustmann, Christian, and Albrecht Glitz. “How Do Industries and Firms Respond to Changes in Local Labor Supply?” *Journal of Labor Economics* 33, 3: (2015) 711–750.
- Dustmann, Christian, Albrecht Glitz, and Thorsten Vogel. “Employment, wages, and the economic cycle: Differences between immigrants and natives.” *European Economic Review* 54, 1: (2010) 1–17.
- Dustmann, Christian, Uta Schönberg, and Jan Stuhler. “Labor Supply Shocks, Native Wages and the Adjustment of Local Employment.” *Quarterly Journal of Economics* 132, September: (2017a) 435–483.
- . “Labor Supply Shocks, Native Wages, and the Adjustment of Local Employment.” *The Quarterly Journal of Economics* 132, 1: (2017b) 435–483.
- Easterly, W. *The Elusive Quest for Growth*. 2001.
- Eberle, Johanna, Peter Jacobebbinghaus, Johannes Ludsteck, and Julia Witter. “Generation of time-consistent industry codes in the face of classification changes : Simple heuristic based on the Establishment History Panel (BHP).” FDZ-Methodenreport 05/2011 (en), Institute for Employment Research (IAB), Nürnberg, 2011. https://ideas.repec.org/p/iab/iabfme/201105{}_en.html.
- Edo, Anthony. “How do rigid labor markets absorb immigration? Evidence from France.” *IZA Journal of Migration* 5, 1: (2016) 7.
- . “Evidence from the Algerian Independence War.” *Cepii Discussion Series* .
- Elsner, Benjamin. “Does emigration benefit the stayers? Evidence from EU enlargement.” *Journal of Population Economics* 26, 2: (2013) 531–553.
- Fally, Thibault, Rodrigo Paillacar, and Cristina Terra. “Economic geography and wages in Brazil: Evidence from micro-data.” *Journal of Development Economics* 91, 1: (2010) 155–168.
- Farré, Lidia, Libertad González, and Francesc Ortega. “Immigration , Family Responsibilities and the Labor Supply of Skilled Native Women.” *B . E . Journal of Economic Analysis & Policy* 11, 1.
- Feenstra, Robert C., Robert E. Lipsey, Haiyan Deng, Alyson C. Ma, and Hengyong Mo. “World Trade Flows: 1962-2000.”, 2005.
- Felbermayr, Gabriel J., Sanne Hiller, and Davide Sala. “Does immigration boost per capita income?” *Economics Letters* 107, 2: (2010) 177–179.
- Fernandez, Cristina, and Carolina Ortega. “Labor market assimilation of immigrants in Spain: employment at the expense of bad job-matches?” *Spanish Economic Review* 10: (2008) 83–107.
- Fernández-Huertas Moraga, Jesús. “Immigrant Selection over the Business Cycle The Spanish Boom and the Great Recession.” *Documentos de trabajo (FEDEA)* 5: (2014) 1–36.

- Fernández-Huertas Moraga, Jesus. “Immigrant Locations and Native Residential Preferences.”, 2017.
- Fernández-Huertas Moraga, Jesús, Ada Ferrer-i Carbonell, and Albert Saiz. “Immigrant Locations and Native Residential Preferences : Emerging Ghettos or New Communities ?” *IZA Discussion Papers* , 11143.
- Finger, JM, and ME Kreinin. “A Measure of ‘Export Similarity’ and Its Possible Uses.” *The Economic Journal* 89, 356: (1979) 905–912.
- Fitzenberger, Bernd, Aderonke Osikominu, and Robert Völter. “Imputation Rules to Improve the Education Variable in the IAB Employment Subsample.” *Schmollers Jahrbuch : Journal of Applied Social Science Studies* 126, 3: (2006) 405–436.
- Font, Paulino, Mario Izquierdo, and Sergio Puente. “Real wage responsiveness to unemployment in Spain : asymmetries along the business cycle.” *IZA Journal of European Labor Studies* 4, 13. ???
- Foster, Neil, and Robert Stehrer. “Sectoral productivity, density and agglomeration in the wider Europe.” *Spatial Economic Analysis* 4, 4: (2009) 427–446.
- Frankel, By Jeffrey A, and David Romer. “Does Trade Cause Growth?” *American Economic Review* 89, 3: (2019) 379–399.
- Friedberg, Rachel M. “The Impact of Mass Migration on the Israeli Labor Market.” *Quarterly Journal of Economics* 116, 4: (2001) 1373–1408.
- Gezici, Ferhan, and Geoffrey J.D. Hewings. “Regional convergence and the economic performance of peripheral areas in Turkey.” *Review of Urban and Regional Development Studies* 16, 2: (2004) 113–132.
- Gil-Alonso, Fernando, Jordi Bayona-I-Carrasco, and Isabel Pujadas Rúbies. “Foreigner’s Internal Migration in Spain: Recent Spatial Changes During the Economic Crisis.” 69–2015.
- Giorcelli, Michela. “No The Long-Term Effects of Management and Technology Transfers.” *American Economic Review* 109, 1: (2019) 121–152.
- Glaeser, Edward, and J. Vernon Henderson. “Urban economics for the developing World: An introduction.” *Journal of Urban Economics* 98: (2017) 1–5.
- Glaeser, Edward L. “Learning in Cities.” *Journal of Urban Economics* 46, 2: (1999) 254–277.
- Glaeser, Edward L., and Joshua D. Gottlieb. “The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States.” *Journal of Economic Literature* 47, 4: (2009) 983–1028.

- Glaeser, Edward L., and David C. Maré. “Cities and Skills.” *Journal of Labor Economics* 19, 2: (2001) 316–342.
- Glitz, Albrecht. “The Labor Market Impact of Immigration: A Quasi-Experiment Exploiting Immigrant Location Rules in Germany.” *Journal of Labor Economics* 30, 1: (2012) 175–213.
- Glitz, Albrecht, and Erik Meyersson. “Industrial Espionage and Productivity.” 25–27.
- Goldin, Claudia, and Lawrence Katz. “The Race between Education and Technology: The Evolution of U.S. Educational Wage Differentials, 1890 to 2005.” Technical report, National Bureau of Economic Research, Cambridge, MA, 2007.
- Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. “Bartik Instruments: What, When, Why, and How.” Technical report, 2018.
- Gonzalez, Libertad, and Francesc Ortega. “How do very open economies adjust to large immigration flows? Evidence from Spanish regions.” *Labour Economics* 18, 1: (2011) 57–70.
- . “Immigration and housing booms: Evidence from Spain*.” *Journal of Regional Science* 53, 1: (2013) 37–59.
- Gonzalez, Libertad, Francesc Ortega, and Lúcia Farré. “Immigration, Family Responsibilities and the Labor Supply of Skilled Native Women.”, 2009.
- González-Ferrer, Amparo, and Francisco Javier Moreno-Fuentes. “Back to the Suitcase? Emigration during the Great Recession in Spain.” *South European Society and Politics* 22, 4: (2017) 447–471.
- Gould, DM. “Immigrant links to the home country: empirical implications for US bilateral trade flows.” *The Review of Economics and Statistics* 76, 2: (1994) 302–316.
- Groot, Stefan P.T., Henri L.F. de Groot, and Martijn J. Smit. “Regional wage differences in the Netherlands: Micro evidence on agglomeration externalities.” *Journal of Regional Science* 54, 3: (2014) 503–523.
- Gutiérrez-Portilla, María, Adolfo Maza, and María Hierro. “Foreigners versus natives in Spain: different migration patterns? Any changes in the aftermath of the crisis?” *Annals of Regional Science* 61, 1: (2018a) 139–159.
- Gutiérrez-Portilla, María, José Villaverde, Adolfo Maza, and María Hierro. “A spatial approach to the impact of immigration on wages: evidence from Spain.” *Regional Studies* 3404: (2018b) 1–10.
- Hailbronner, Kay. “National Report Germany.” In *Study on the Single Asylum Procedure One-stop-shop against the Background of the Common European Asylum System and the Goal of a Common Asylum Procedure*, edited by Kay Hailbronner, Luxemburg, 2003, 221–244.

- Harris, Chauncy D. “The market as a factor in the location of industry in the United States.” *Annals of the Association of American Geographers* 44, 4: (1954) 315–348.
- Hausmann, Ricardo, César A Hidalgo, Sebastián Bustos, Michele Coscia, Alexander Simoes, and Muhammed A. Yildirim. *The Atlas of Economic Complexity: Mapping Paths to Prosperity*. Cambridge, MA: MIT Press, 2014.
- Hausmann, Ricardo, and Frank Neffke. “The workforce of pioneer plants.”, 2016.
- Hausmann, Ricardo, Francisco Rodríguez, and Rodrigo Wagner. “Growth Collapses.” *CID Working Paper Series* , RWP06-046.
- Head, Keith, Thierry Mayer, and John Ries. “The erosion of colonial trade linkages after independence.” *Journal of International Economics* 81, 1: (2010) 1–14.
- Henderson, Vernon, Todd Lee, and Yung Joon Lee. “Scale externalities in Korea.” *Journal of Urban Economics* 49, 3: (2001) 479–504.
- Hering, Laura, and Sandra Poncet. “Market access and individual wages: Evidence from China.” *Review of Economics and Statistics* 92, 1: (2010) 145–159.
- Hidalgo, César A, Bailey Klinger, AL Barabási, and Ricardo Hausmann. “The product space conditions the development of nations.” *Science* 317, 5837: (2007) 482–7.
- Hierro, María, and Adolfo Maza. “Per capita income convergence and internal migration in Spain: Are foreign-born migrants playing an important role?” *Papers in Regional Science* 89, 1: (2010) 89–107.
- Hornung, Erik. “Immigration and the Diffusion of Technology : The Huguenot Diaspora in Prussia.” *American Economic Review* 104, 1: (2014) 84–122.
- Hunt, Jennifer. “The Impact of the 1962 Repatriates from Algeria on the French Labor Market.” *Industrial and Labor Relations Review* 45, 3: (1992) 556.
- . “The Impact of Immigration on the Educational Attainment of Natives.” *Journal of Human Resources* 52, 4: (2017) 1060–1118.
- IAB employment history (BeH). “No Title.”, 2015.
- Imbs, Jean, and Romain Wacziarg. “Stages of Diversification.” *American Economic Review* 93, 1: (2003) 63–86.
- Ismihan, Mustafa, and Kivilcim Metin-Ozcan. “Productivity and Growth in an Unstable Emerging Market Economy: The Case of Turkey, 1960-2004.” *Emerging Markets Finance and Trade* 45, 5: (2009) 4–18.

- Izquierdo, Mario, Juan F. Jimeno, and Aitor Lacuesta. "Spain: from massive immigration to vast emigration?" *IZA Journal of Migration* 5, 1: (2016) 1–20.
- Izquierdo, Mario, and Juan Francisco Jimeno. "Employment, wage and price reactions to the crisis in Spain: firm-level evidence from the WDN Survey." *Documentos Ocasionales* , 1503.
- Izquierdo, Mario, Aitor Lacuesta, and Raquel Vegas. "Assimilation of immigrants in Spain: A longitudinal analysis." *Labour Economics* 16, 6: (2009) 669–678.
- Jacobs, Jane. *The Economy of Cities*. New York: Vintage Books, 1969.
- Jaeger, David A, Joakim Ruist, and Jan Stuhler. "Shift-Share Instruments and the Impact of Immigration.", 2018.
- Jauer, Julia, Thomas Liebig, John P. Martin, and Patrick A. Puhani. "Migration as an adjustment mechanism in the crisis? A comparison of Europe and the United States 2006 and 2016." *Journal of Population Economics* 32, 1: (2019) 1–22.
- Javorcik, Beata S., Caglar Ozden, Mariana Spatareanu, and Cristina Neagu. "Migrant networks and foreign direct investment." *Journal of Development Economics* 94, 2: (2011) 231–241.
- Karpat, Kemal H. *Ottoman Population 1830-1914*. Madison, WI: University of Wisconsin Press, 1985.
- Kemeny, Thomas, and Abigail Cooke. "Spillovers from immigrant diversity in cities." *Journal of Economic Geography* 18, 1: (2018) 213–245.
- Kerr, Sari Pekkala, and William R. Kerr. "Global Collaborative Patents." *Economic Journal* 128: (2018) 235–272.
- Kerr, W R. "Ethnic Scientific Communities and International Technology Diffusion." *Review of Economics and Statistics* 90, 3: (2008) 518–537.
- Kerr, William. *The Gift of Global Talent: How Migration Shapes Business, Economy & Society*. Stanford: Stanford University Press, 2018.
- Khanna, Gaurav, and Nicolas Morales. "The it Boom and Other Unintended Consequences of Chasing the American Dream.", 2017.
- Kirdar, Murat G., and Sirin D. Saracoglu. "Migration and regional convergence: An empirical investigation for Turkey." *Papers in Regional Science* 87, 4: (2008) 545–566.
- Koren, Miklos, and Silvana Tenreyro. "Volatility and Development." *The Quarterly Journal of Economics* 122, 1: (2007) 243–287.

- Krishna, Pravin, and Andrei A. Levchenko. “Comparative advantage, complexity, and volatility.” *Journal of Economic Behavior and Organization* 94: (2013) 314–329.
- Krueger, Anne O., and Baran Tuncer. “Growth of factor productivity in Turkish manufacturing industries.” *Journal of Development Economics* 11, 3: (1982) 307–325.
- Krugman, P., and A. J. Venables. “Globalization and the Inequality of Nations.” *The Quarterly Journal of Economics* 110, 4: (2006) 857–880.
- Krugman, Paul R. “Scale Economies, Product Differentiation, and the Pattern of Trade.” *American Economic Review* 70, 5: (1980) 950–959.
- Kugler, Maurice, Oren Levintal, and Hillel Rapoport. “Migration and Cross-Border Financial Flows Migration and Cross-Border Financial Flows.” *World Bank Economic Review* .
- Kugler, Maurice, and Hillel Rapoport. “International labor and capital flows: Complements or substitutes?” *Economics Letters* 94, 2: (2007) 155–162.
- Lafuente, Cristina. “Unemployment in administrative data using survey data as a benchmark.” *SERIEs* .
- Lall, Somik V., Zmarak Shalizi, and Uwe Deichmann. “Agglomeration economies and productivity in Indian industry.” *Journal of Development Economics* .
- Lederer, Harald. *Migration und Integration in Zahlen*. Bonn, 1997, beauftragt edition.
- Lee, Bun Song, Soomyung Jang, and Sung Hyo Hong. “Marshall’s scale economies and Jacobs’ externality in Korea: The role of age, size and the legal form of organisation of establishments.” *Urban Studies* 47, 14: (2010) 3131–3156.
- Lee, Jongkwan, Giovanni Peri, and Vasil Yassenov. “The Employment Effects of Mexican Repatriations: Evidence from the 1930’s.” Technical report, 2017.
- Leromain, Elsa, and Gianluca Orefice. “New revealed comparative advantage index: Dataset and empirical distribution.” *International Economics* 139: (2014) 48–70.
- Lewis, Ethan, and Giovanni Peri. *Immigration and the Economy of Cities and Regions*, volume 5. Elsevier B.V., 2015, 1 edition.
- Liedtke, Matthias. “National welfare and asylum in Germany.” *Critical Social Policy* 22, 3: (2002) 479–497.
- López, Ricardo A., and Jens Südekum. “Vertical industry relations, spillovers, and productivity: Evidence from Chilean plants.” *Journal of Regional Science* 49, 4: (2009) 721–747.

- Luca, Davide. “Votes and Regional Economic Growth: Evidence from Turkey.” *World Development* 78: (2016) 477–495.
- Lucas, Robert E. “On the mechanics of economic development.” *Journal of Monetary Economics* 22, 1: (1988) 3–42.
- MacKinnon, JG, and Lonnie Magee. “Transforming the Dependent Variable in Regression Models.” *International Economic Review* 31, 2: (1990) 315–339.
- Malchow-Møller, Nikolaj, Jakob Roland Munch, and Jan Rose Skaksen. “Do Foreign Experts Increase the Productivity of Domestic Firms?” *The Scandinavian Journal of Economics* 12, 10: (2017) 3218–3221.
- Malo, Miguel A. “Labour Market Measures in Spain 2008 2013 : The Crisis and Beyond Author.” Technical report, International Labour Office, Research Department, Geneva, 2015.
- Markusen, James R., and Natalia Trofimenko. “Teaching locals new tricks: Foreign experts as a channel of knowledge transfers.” *Journal of Development Economics* 88, 1: (2009) 120–131.
- Marshall, Alfred. *Principles of Economics*. London: Macmillan, 1890.
- Martin, Philippe, Thierry Mayer, and Florian Mayneris. “Spatial concentration and plant-level productivity in France.” *Journal of Urban Economics* 69, 2: (2011) 182–195.
- Maza, Adolfo, María Gutiérrez-Portilla, María Hierro, and José Villaverde. “Internal Migration in Spain: Dealing with Multilateral Resistance and Nonlinearities.” *International Migration* 57, 1: (2019) 75–93.
- McCann, Philip. “Perceptions of regional inequality and the geography of discontent: insights from the UK.” *Regional Studies* .
- Melguizo, Celia, and Vicente Royuela. “What drives migration moves across urban areas in Spain? Evidence from the Great Recession.” *SSRN* .
- Melo, Patricia C., Daniel J. Graham, and Robert B. Noland. “A meta-analysis of estimates of urban agglomeration economies.” *Regional Science and Urban Economics* 39, 3: (2009) 332–342.
- Messina, Julián, Cláudia Filipa Duarte, Mario Izquierdo, Philip Caju, and Niels Lynggård Hansen. “The Incidence of Nominal and Real Wage Rigidity: An individual-based sectoral approach.” *Journal of the European Economic Association* 8, 2-3: (2010) 487–496.
- Metin, Karadag, A Onder Ozlem, and Deliktas Ertugrul. “Growth of factor productivity in the Turkish manufacturing industry at provincial level.” *Regional Studies* 39, 2: (2005) 213–223.
- Mion, Giordano, and Paolo Naticchioni. “The spatial sorting and matching of skills and firms.” *Canadian Journal of Economics* 42, 1: (2009) 28–55.

- Mishra, Prachi. “Emigration and wages in source countries: Evidence from Mexico.” *Journal of Development Economics* 82, 1: (2007) 180–199.
- Monras, Joan. “Economic Shocks and Internal Migration.” *Quantitative Economics* 1–71.
- . “Immigration and Wage Dynamics: Evidence from the Mexican Peso Crisis.” *Journal of Political Economy* (Forthcoming) .
- Monras, Joan, Javier Vázquez-Grenno, and Ferran Elias. “Understanding the Effects of Legalizing Undocumented Immigrants.” *CEPR Discussion Paper 12726* .
- Moral-benito, Enrique. “The Microeconomic origins of the Spanish boom.” *Bank of Spain Working Paper* 1805.
- Moretti, Enrico. “Estimating the Social Return to Education: Evidence from Repeated Cross-Sectional and Longitudinal Data.”, 1999.
- Moulton, Brent R. “An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units.” *The Review of Economics and Statistics* 72, 2: (1990) 334.
- Mulhern, Alan, and John G. Watson. “Spanish internal migration: Is there anything new to say?” *Spatial Economic Analysis* 4, 1: (2009) 103–120.
- Nathan, M. “After Florida: Towards an economics of diversity.” *European Urban and Regional Studies* 22, 1: (2015) 3–19.
- Nenadic, Mario, Nermina Džepar-Ganibegović, Medžid Lipjankić, Drago Borovčanin, Nada Spasojević, Dragana Kovac, and Abela Pobrić. “Comparative Analysis on Access to Rights of Refugees and Displaced Persons.” Ministry for Human Rights and Refugees Bosnia and Herzegovina, 2005.
- Nunn, Nathan. “Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade.” *The Quarterly Journal of Economics* 122, 2: (2007) 569–600.
- OECD. *International Migration and the Economic Crisis*. Paris, 2009.
- . *OECD Regions and Cities at a Glance 2018*. Paris, 2018.
- . *International Migration Outlook 2019*. Paris, 2019.
- Onder, A, E Deliktas, and A Lenger. “Efficiency in the Manufacturing Industry of Selected Provinces in Turkey: A Stochastic Frontier Analysis.” *Emerging Markets Finance and Trade* 39, 2: (2003) 98–113.
- Orrenius, Pia M., and Madeline Zavodny. “Does immigration affect wages? A look at occupation-level evidence.” *Labour Economics* 14, 5: (2007) 757–773.

- Orrenius, Pia M, and Madeline Zavodny. “Mexican Immigrant Employment Outcomes over the Business Cycle.” *American Economic Review: Papers & Proceedings* 100, 2.
- Ortega, Francesc, and Giovanni Peri. “The Causes and Effects of International Migrations :Evidence From OECD Countries 1980-2005.” *NBER Working Paper* .
- . “Openness and income: The roles of trade and migration.” *Journal of International Economics* 92, 2: (2014) 231–251.
- Ortega, Javier, and Gregory Verdugo. “Moving Up or Down? Immigration and the Selection of Natives across Occupations and Locations.” 106–112.
- Ottaviano, Gianmarco I.P., and Giovanni Peri. “Immigration and national wages: Clarifying the theory and the empirics.” *SSRN Electronic Journal* .
- . “Rethinking the effect of immigration on wages.” *Journal of the European Economic Association* 10, 1: (2012) 152–197.
- Ottaviano, Gianmarco I.P. P, and Giovanni Peri. “The economic value of cultural diversity: Evidence from US cities.” *Journal of Economic Geography* 6, 1: (2006) 9–44.
- Özden, ÇaÇşlar, Christopher R. Parsons, Maurice Schiff, and Terrie L. Walmsley. “Where on earth is everybody? The evolution of global bilateral migration 1960-2000.” *World Bank Economic Review* 25, 1: (2011) 12–56.
- Pamuk, Sevket. “Agricultural Output and Productivity Growth in Turkey since 1880.” In *Agriculture and Economic Development in Europe since 1870*, edited by P. Lains, and V Pinilla, London, 2008. Routledge edition.
- Parsons, Christopher, and Pierre-Louis Vézina. “Migrant Networks and Trade: The Vietnamese Boat People as a Natural Experiment.” *The Economic Journal* 128, 612: (2018a) F210–F234.
- Parsons, Christopher R, and Pierre-Louis Louis Vézina. “Migrant Networks and Trade: The Vietnamese Boat People as a Natural Experiment.” *Economic Journal* 128, 612: (2018b) F210–F234. <http://doi.wiley.com/10.1111/eoj.12457><https://onlinelibrary-wiley-com.ezproxy.univ-paris1.fr/doi/pdf/10.1111/eoj.12457>.
- Peri, Giovanni. “The effect of immigration on productivity: Evidence from U.S. states.” *The Review of Economics and Statistics* 94, 1: (2012) 348–358.
- . “Immigrants, Productivity, and Labor Markets.” *Journal of Economic Perspectives* 30, 4: (2016) 3–30.

- Peri, Giovanni, and Chad Sparber. “Assessing inherent model bias: An application to native displacement in response to immigration.” *Journal of Urban Economics* 1, 1: (2011) 82–91.
- Peri, Giovanni, Chad Sparber, By Giovanni Peri, and Chad Sparber. “Task Specialization, Immigration, and Wages.” *American Economic Journal: Applied Economics* 1, 3: (2009) 135–169.
- Polanyi, M. *The Tacit Dimension*. Chicago; London: University of Chicago Press, 1966, 2009 edition.
- Pomeranz, Kenneth. *The Great Divergence: China, Europe, and the Making of the Modern World Economy*. Princeton University Press, 2001.
- Poole, Jennifer. “Knowledge Transfers from Multinational to Domestic Firms: Evidence from Worker Mobility.” *The Review of Economics and Statistics* 95, 2: (2013a) 393–406.
- Poole, Jennifer P. “Knowledge transfers from multinational to domestic firms: Evidence from worker mobility.” *Review of Economics and Statistics* 95, 2: (2013b) 393–406. http://www.mitpressjournals.org/doi/10.1162/REST_{_}a_{_}00258.
- Puente, Sergio, and Sofía Galán. “Un análisis de los efectos composición sobre la evolución de los salarios.” *Boletín Económico Banco de E*, 2006: (2014) 57–62.
- Radović, B, G Opačić, and I Vidaković. “A Brief Retrospective on the Problem of Refugees in the Yugoslav Wars 1991-99.” *Living in post-war* .
- Rauch, James E. “Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities.” *Journal of Urban Economics* 34, 3: (1993) 380–400.
- Rauch, James E. “Networks versus markets in international trade.” *Journal of International Economics* 48, 1: (1999) 7–35.
- Rauch, JE, and Vitor Trindade. “Ethnic Chinese networks in international trade.” *Review of Economics and Statistics* 84, February: (2002) 116–130.
- Redding, S, and AJ Venables. “Redding and Venables 2003 Economic Geography and International Inequality.” *Journal of international Economics* .
- Redding, Stephen J., and Matthew A. Turner. “Transportation Costs and the Spatial Organization of Economic Activity.” In *Handbook of Regional and Urban Economics*, Elsevier, 2015, volume 5, 1339–1398.
- Rhee, Yung, and Therese Belot. “Export Catalysts in Low-Income Countries.”, 1990.
- de la Rica, Sara, Albrecht Glitz, and Francesc Ortega. “Immigration in Spain: what have we learned from recent evidence?” *Cuadernos económicos de ICE* 87: (2014) 9–28.

- Rice, Patricia, Anthony J. Venables, and Eleonora Patacchini. "Spatial determinants of productivity: Analysis for the regions of Great Britain." *Regional Science and Urban Economics* 36, 6: (2006) 727–752.
- Roback, Jennifer. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90, 6: (1982) 1257–1278.
- Ródenas, Carmen, Mónica Martí, and Ángel León. "A new pattern in international mobility? The case of Spain in the Great Crisis." *Investigación Económica* 76, 299: (2017) 153–181.
- Rodenas Calatayud, Carmen, and Monica Marti Sempere. "L'estimation des fausses migrations en Espagne." *Population* 64, 2: (2009) 397–412.
- Rodríguez-Pose, Andrés. "The revenge of the places that don't matter (and what to do about it)." *Cambridge Journal of Regions, Economy and Society* 11, 1: (2018) 189–209.
- Rosenberg, David. *Israel's Technology Economy*. Cham: Palgrave Macmillan, 2018.
- Rosenthal, Stuart S., and William C. Strange. "Evidence on the nature and sources of agglomeration economies." In *Handbook of Regional and Urban Economics*, Elsevier Inc., 2004, volume 4, 2119–2171.
- . "The attenuation of human capital spillovers." *Journal of Urban Economics* 64, 2: (2008) 373–389.
- Rothstein, Jesse. "The labor market four years into the crisis: Assessing structural explanations.", 2012.
- Ruben, Ruerd, Marieke Van Houte, and Tine Davids. "What determines the embeddedness of forced-return migrants? Rethinking the role of pre- and post-return assistance." *International Migration Review* 43, 4: (2009) 908–937.
- Ruhl, Stefan, and Harald W Lederer. *Migrationsbericht 2001 der Ausländerbeauftragten im Auftrag der Bundesregierung*. Beauftragte der Bundesregierung für Ausländerfragen, 2001.
- Rybczynski, T. M. "Factor Endowment and Relative Commodity Prices." *Economica* 22, 88: (1955) 336–341.
- Saito, Hisamitsu, and Munisamy Gopinath. "Plants' self-selection, agglomeration economies and regional productivity in Chile." *Journal of Economic Geography* 9, 4: (2009) 539–558.
- Saiz, Albert. "Immigration and housing rents in American cities." *Journal of Urban Economics* 61, 2: (2007) 345–371.
- Sakalli, Seyhun Orcan. "Secularization and Religious Backlash : Evidence from Turkey.", 2019.
- Sanchis-Guarner, Rosa. "Decomposing the Impact of Immigration on House Prices." *Ssrn* , September.

- Schmucker, Alexandra, Johanna Eberle, Andreas Ganzer, Jens Stegmaier, and Matthias Umkehrer. “Establishment History Panel 1975-2016.” Fdz-datenreport, 01/2018 (en), Institut für Arbeitsmarkt- und Berufsforschung (IAB), 2018.
- Schündeln, Matthias. “Are immigrants more mobile than natives? Evidence from Germany.” *Journal of Regional Science* 54, 1: (2014) 70–95.
- Shirotori, M, B Tumurchudur, and O Cadot. “Revealed Factor Intensity Indices at the Product Level.” *Policy Issues in International Trade and Commodities* 2010, 44.
- Solon, Gary, Steven J Haider, and Jeffrey Wooldridge. “What Are We Weighting For.” *The Journal of Human Resources* 50, 2: (2015) 301–316.
- Spitz-Oener, Alexandra. “Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure.” *Journal of Labor Economics* 24, 2: (2006) 235–270.
- Syverson, Chad. “Challenges to mismeasurement explanations for the US productivity slowdown.” *Journal of Economic Perspectives* 31, 2: (2017) 165–186.
- Temel, Tugrul, Aysit Tansel, and Peter J. Albersen. “Convergence and Spatial Patterns in Labor Productivity: Nonparametric Estimation for Turkey.”, 1999.
- Trax, Michaela, Stephan Brunow, and Jens Suedekum. “Cultural diversity and plant-level productivity.” *Regional Science and Urban Economics* 53, 119: (2015) 85–96.
- Tumen, Semih. “The Economic Impact of Syrian Refugees on Host Countries: Quasi-Experimental Evidence from Turkey.” *American Economic Review* 106, 5: (2016) 456–460.
- UNHCR. “After the war was over.” *Refugee magazine* <http://www.unhcr.org/433bded34.pdf>.
- Verdugo, Gregory. “Real wage cyclicalities in the Eurozone before and during the Great Recession: Evidence from micro data.” *European Economic Review* 82: (2016) 46–69.
- Walker, William S. *German and Bosnian voices in a time of crisis: Bosnian refugees in Germany, 1992-2002*. Indianapolis: Dog Ear Publishing, 2010.
- World Bank. “Rise of the Anatolian Tigers.” Technical Report 87180, The World Bank Group, Washington DC, 2015.
- Yagan, Danny. “Is the Great Recession Really Over? Longitudinal Evidence of Enduring Employment Impacts.” *Journal of Political Economy* (Forthcoming) .

List of Figures

1	Income and productivity	10
2	Foreign-born population	13
3	Foreign-born population, by year and country	14
4	Revenu par habitant et productivité du travail	24
5	Population migrant	27
6	Population migrant, par année et pays	29
7	Annual Unemployment Rate 2004-2017	48
8	Immigrant Population 1998-2017	50
9	International Flows of immigrant	53
10	Yearly Net Change in Immigrant Population	54
11	Population Responses to Employment Shocks: Native-Born and Immigrant Men	55
12	Net Change in immigrant Population by Province 2009-2014	61
13	Unemployment and Out-migration	65
14	Instrumental Variable Relevance	69
15	Unemployment Rates	103
16	Native vs immigrant Men	105
17	International Flows by Nationality	109
18	Domestic vs. international departures by Nationality	111
19	Immigrant Population by Skills	112
20	Population by Average Age	113
21	Net Change in immigrant population by province 2009-2016	115
22	Immigrant and native outflows	125
23	Maps	126
24	Migration from Yugoslavia into Germany	139
25	Yugoslavian workers yearly entry to and exit from German's labor force	140
26	Yugoslavians in German workforce, by product	142
27	Exports for products with different levels of treatment	143
28	Distribution rule of asylum seekers in Germany 1995	146
29	Instrumental Variable Relevance	147
30	Share of arrival vs exit from the German labor force by product	149
31	Event study (2SLS), 5 year periods	156
32	“Placebo” test using exports from similar countries	159
33	“Placebo” test using return migrants to other countries	160
34	Treatment effect for industries with high share of early returnees	162

35	Marginal effect by type of migrant	174
36	Exploration of zeros in our data	182
37	Spatial Differences	213
38	Instrumental Variable Relevance	230

List of Tables

1	Summary Statistics	62
2	First Stage Regressions, 2009-2014	70
3	Effects on Wages and Employment of Natives, 2009-2014	71
4	Falsification Test 1, Growth in 2003-2008	72
5	Falsification Test 2 , Pre-trends	73
6	Falsification Test 3 , Share of immigrant	73
7	Wages and Employment with Controls	75
8	Alternative Instruments	78
9	Alternative Measures	80
10	Alternative Measures	81
11	Alternative Weights	82
12	Longer Differences: Employment	83
13	Wages and Employment by Skill Group	84
14	Wages and Employment by Demographic Groups	85
15	EU vs. Non-EU outflows	86
16	Impact of out-migration on native mobility between 2009-2014	88
17	Impact of out-migration on native mobility between 2009-2014	89
18	Entries versus Exits: Employment	91
19	New Entries vs Already Employed: Wages	92
20	Daily Wages vs Days Worked in a month	92
21	Wages and Employment: Contract Types	94
22	Foreign-Mobility Smooths Employment Outcomes	106
23	immigrant Outflows Smooths Employment Outcomes	107
24	Migrant Stocks EU vs. non-EU	110
25	Robustness: Three Year Average Treatment, 2009-2014	116
26	Gravity Equation: Predicting National Stocks	120
27	IV validity checks: Base-year validity regressions	122
28	Multiple Instrumentation, 2009-2016	123
29	Boom vs. Bust	127
30	Part-time Employment and Wages	129
31	Summary Statistics Yugoslavian Refugees in Germany	150
32	Difference-in-difference estimation	151
33	DID estimation, previous trend	154
34	Event study 5-year periods	155

35	DID, controlling for German FDI	157
36	DID, Costinot et al. (2012) measures in LHS	161
37	DID, outcomes in Bosnia	165
38	DID, by region of destination	167
39	DID, heterogeneity analysis	169
40	DID, workers' education and occupations	172
41	DID, including exports to Germany	184
42	DID, controlling for global FDI	185
43	Exports vs. FDI	186
44	DID, controlling for imports of intermediate goods	187
45	DID, WZ 93 3-digit aggregation	189
46	DID, clustering by W93 industries	190
47	DID, different treatments	191
48	DID, no zeros	191
49	DID, excl. Slovenia	192
50	DID, excl. Macedonia	192
51	Occupations List	193
52	Occupations List (cont.)	194
53	Difference-in-differences (OLS), workers weighted by wage	196
54	Difference-in-differences (OLS), workers' education	197
55	Difference-in-differences (OLS), workers' occupation tasks	198
56	Difference-in-differences (OLS), workers' occupation skill level	199
57	Difference-in-differences (OLS), workers' occupation supervision intensity	200
58	Difference-in-differences (OLS), workers' wage growth	201
59	Summary statistics, all countries	203
60	DID, all countries	204
61	OLS	228
62	First Stage Regressions, 2008-2013	231
63	2SLS Results for 2008-2013	232
64	Explanatory Power of Various Sets of Variables	234
65	Second Stage Regressions, 2008-2013	235
66	2SLS Results for 2008-2013	238
67	OLS Results for 2008-2013: Infrastructure and Amenities	241
68	Summary Statistics for Local Variables	244
69	Summary Statistics for Individual Anaalysis	244

70	Yearly Regressions	245
71	One Step Regressions, 2008-2013	246
72	First Step Regressions, 2008-2013	247
73	Second Stage Regressions, 2008-2013	248
74	Accounting for informal sector, 2008-2013	250
75	First Stage Regressions, 2008-2013	252
76	OLS Results for 2008-2013: Infrastructure and Amenities	253

Essais sur les migrations et la productivité

Cette thèse s'intéresse à l'interaction entre migration et productivité, sous différents angles, à travers trois pays et périodes. Plus précisément, j'étudie les avantages à la mobilité pour les migrants sur le marché du travail en période de crise économique, les gains de productivité dus à la mobilité des migrants lors de la reconstruction d'un pays au lendemain d'une guerre, et les gains associés à une plus forte concentration de la population dans les grandes zones urbaines. J'aborde ces sujets à la fois théoriquement et empiriquement, en utilisant de riches données confidentielles sur la sécurité sociale provenant d'Espagne, d'Allemagne et de Turquie, en utilisant diverses techniques de données de panel ainsi que des instruments historiques pour estimer des relations causales. Les résultats de ces études portent sur de nombreuses questions qui intéressent à la fois le monde universitaire et les décideurs politiques, mais sur lesquelles on sait encore peu de choses. Cette thèse vise à contribuer à améliorer nos connaissances sur des questions qui demeureront pertinentes dans un avenir proche.

Mots-clés: migration, réfugiés, diffusion des connaissances, exportations, productivité, marchés du travail locaux, salaires, emploi, Turquie, Espagne, Allemagne

Essays on Migration and Productivity

This dissertation explores the interaction between migration and productivity, through multiple angles, across three different country and period contexts. Specifically, I study the labor market benefits of migrant mobility during an economic crisis, productivity gains due to migrant mobility in the reconstruction of a country in the aftermath of a war, and gains associated with a higher concentration of people in larger urban areas. I address these subjects both theoretically and empirically, using rich confidential social security data from Spain, Germany, and Turkey, applying a variety of panel data techniques and historical instruments to estimate causal relationships. The findings of these studies relate to many issues that interest both the academia and the policymakers yet on which little is known. This dissertation aims to contribute to knowledge gap on issues that will remain relevant foreseeable future.

Keywords: migration, refugees, knowledge diffusion, exports, productivity, local labour markets, wages, employment, Turkey, Spain, Germany