CERGE
Center for Economic Research and Graduate Education
Charles University

Essays on Citizenship Policies and Immigrant Integration

Vahan Sargsyan

Dissertation

Prague, October 2018

Dissertation Committee:

Jan Hanousek (CERGE-EI, Charles University, Chair)

Mariola Pytliková (CERGE-EI, Charles University)

Nikolas Mittag (CERGE-EI, Charles University)

Alena Bičáková (CERGE-EI, Charles University)

Randall K. Filer (Hunter College, CUNY)

Referees:

Benjamin Elsner (University College Dublin)

Martin Guzi (Masaryk University)

Essays on Citizenship Policies and Immigrant Integration

CERGE
Center for Economic Research and Graduate Education

Charles University Prague

Vahan Sargsyan

Contents

Acknowledgements

I express my gratitude to everyone who has helped me in my dissertation work and throughout my studies. I am very grateful to my supervisor, Jan Hanousek, who has provided me with his continuous guidance and support. The encouragement that I received from him was one of the main reasons for the creation and finalization of this work.

I would also thank all members of my dissertation committee, Mariola Pytliková, Randall K. Filer, Nikolas Mittag and Alena Bičáková, for their advice and comments which helped me shape and improve my work. I would especially like to highlight Mariola Pytliková, who worked with me intensively on the adjustments and improvements of the dissertation after the first submission of the manuscript. Her careful guidance and supervision helped me completely change and formulate the third chapter of this dissertation.

I would like to thank the Academic Skills Center of CERGE-EI, and in particular Deborah Nováková, Andrea Downing, Gray Krueger, and Robin-Eliece Mercury, for teaching me how to write academically and for editing my work.

Special thanks go also to fellow students at CERGE-EI, especially Suren Vardanyan, Liyousew Borga, and Danijela Vuletić for inspiring discussions and advice, and also for their encouragement that kept me motivated during my studies.

Most of all, I would like to thank my family and my wife for their unconditional support, love, and patience.

Financial support from the Czech Science Foundation (grant number 14-36154G), the Grant Agency of Charles University (grant number 490617), and the Grant Agency of Charles University (grant number 310916) is gratefully acknowledged.

All errors remaining in this text are the responsibility of the author.

Abstract

Differential treatment towards minority groups in host societies and labor markets may be a result of both a governmental registration system that fosters unequal rights based on the origins of individuals and the disadvantageous attitude of local employers and the general population towards non-locals.

In the first chapter, I test for differential treatment in the Chinese labor market towards rural migrants with and without urban registration, using data from the Rural to Urban Migration Survey in China. The findings indicate that despite its often-assumed large impact on the differential treatment towards rural migrants, the type of household registration (*hukou*) is not entirely responsible for the local-migrant differences in the total hourly earnings that are not attributable to personal characteristics. The results suggest that even the complete abolishment of the *hukou* system may at most eliminate only a portion of the disadvantageous treatment towards rural female migrants that is not attributable to differences in personal characteristics, and may even have no measurable impact on rural male migrants working in the paid-employment sector in Chinese urban labor markets.

In the second chapter, I conduct an empirical study in order to estimate the impact of naturalization on the labor market integration of first generation immigrants in two European countries, France and Denmark. This chapter contributes to the existing literature by: (1) comparing the employment opportunities and incomes of naturalized and non-naturalized migrants in European labor markets to those of the native population, and (2) attempting to estimate the impact of characteristics of a country's citizenship policy on this relationship. The results suggest the existence of high naturalization premiums and high socioeconomic integration of naturalized migrants in France, a country with relatively soft naturalization policies, but not for Denmark, which has strict naturalization policies.

Finally, in the third chapter, I conduct a descriptive empirical study on three waves of the European Social Survey (2010, 2012, and 2014), in order to estimate the relationships between the attitude of the native population towards immigrants and immigration (ATII) in 20 European countries and the level of social integration and perceived discrimination of first and second

generation immigrants in those countries. The stringency of naturalization policies in the host countries is also taken into consideration as a mechanism of the development of this relationship.

The results confirm that a country's naturalization policies fairly represent most ATII indicators, and that immigrants feel less discriminated against in more welcoming societies. However, no systematic relationship was revealed between the attitude of the native population and the social integration of immigrants. On the other hand, the results suggest lower perceived discrimination and higher social integration of first generation immigrants in countries where the naturalization status of immigrants is more secured, and the possibility of dual nationality is more restricted.

The results also show that the perceived discrimination of immigrants does not decline with the duration of residence in the host countries, but the latter revealed a positive relationship with the social integration of immigrants.

Abstrakt

Rozdílné zacházení se skupinami menšin v hostitelských společnostech a na trhu práce může být jak výsledkem vládního registračního systému, který podporuje nerovná práva na základě původu jednotlivců, tak také znevýhodňujícího postoje místních zaměstnavatelů a všeobecné populace vůči cizím obyvatelům.

V první kapitole testuji přítomnost rozdílného zacházení na čínském trhu práce vůči venkovským migrantům s a bez městské registrace s využitím dat z průzkumu migrace z venkova do měst v Číně. Zjištění ukazují, že i přes často předpokládaný velký dopad na rozdílné zacházení vůči venkovským migrantům, typ registrace domácností (hukou) není zcela zodpovědný za rozdíly v celkových hodinových příjmech u lokálních migrantů, které nelze přičíst osobním charakteristikám. Výsledky naznačují, že i plné odstranění hukou systému může maximálně eliminovat pouze část znevýhodňujícího přístupu vůči venkovským migrujícím ženám, které nejsou způsobeny rozdíly v osobních charakteristikách, a nemusí mít dokonce žádný měřitelný vliv na venkovské mužské migranty pracující v sektoru výdělečných činností na čínských městských trzích práce.

Druhá kapitola empiricky odhaduje vliv naturalizace – udělení občanství – na integraci do pracovního trhu mezi imigranty první generace ve Francii a Dánsku. Přispíváme k současnému poznání ve dvou aspektech. Zaprvé porovnáváme pracovní možnosti a příjmy mezi naturalizovanými, nenaturalizovanými a domorodými pracovníky. Zadruhé studujeme, jak charakteristiky udělování občanství ovlivňují tyto rozdíly. Naše výsledky ukazují vysokou prémii naturalizace a plné socioekonomické integrace naturalizovaných imigrantů ve Francii, zemi s laxními pravidly pro získání občanství. Tato prémie neexistuje v Dánsku, zemi se striktními pravidly pro získání občanství.

Ve třetí kapitole se zaměřuji na vztahy tří charakteristik, jedná se o postoj původní populace vůči imigrantům a imigrací (anglická zkratka ATII), úroveň sociální integrace a vnímání diskriminace první a druhé generace imigrantů v 20 evropských zemích. Přísnost politiky naturalizace v hostitelských zemích je také brána v potaz jako mechanismus rozvoje uvedených vztahů.

Výsledky potvrzují, že politika naturalizace uspokojivě reprezentuje většinu ATII indikátorů jednotlivých zemí a imigranti se cítí být méně diskriminováni ve společnosti, kde jsou více vítáni. Nicméně, nebyl objeven žádný systematický vztah mezi postoji původní populace a sociální integrací imigrantů. Výsledky na druhou stranu naznačují vnímání diskriminace v menším rozsahu a lepší sociální integraci první vlny imigrantů v zemích, kde je lépe zabezpečen proces naturalizace a možnost dvojího občanství je více omezena.

Výsledky také ukazují, že rozsah vnímané diskriminace imigrantů neklesá s dobou pobytu v hostitelských zemích. Naopak sociální integrace imigrantů vykazuje pozitivní vztah s dobou pobytu v hostitelských zemích.

Introduction

Differential treatment faced by minority groups in labor markets and host societies is an important issue in global economics, with immigrant groups (both internal and external) at destinations being the minorities most studied by empirical researchers. The importance of such research is in its potential application to governmental registration systems and naturalization policy evaluations, a subject undergoing intense study in the era of globalization.

In recent decades, studies of the determinants, causes, and benefits of proper governmental registration of immigrants have created a well-developed body of research for many countries. Some studies focus on the impact of various individual characteristics, as well as origin and destination country characteristics, on obtaining governmental registration in destination countries. Others attempt to estimate the socio-economic benefits of governmental registration via testing for the impact of destination country citizenship on employment probabilities and employment wages. However, it is not yet completely clear whether differential treatment towards minority groups is solely a result of a governmental registration system that fosters unequal rights based on the origins of individuals, or whether it is a result of a negative attitude of both local employers and the general population towards non-locals.

In this dissertation, I aim to create a more comprehensive picture of immigrant integration in host societies and labor markets by first studying the importance of governmental registration in the successful socio-economic integration of immigrants, and then attempting to link the opinion of the native population and institutionally imposed restrictions in the form of governmental policies with the social integration of immigrants and the social acceptance of immigrants in host societies.

The socio-economic integration of immigrants and their children in destination countries is important for many reasons. Most importantly, more successful integration of immigrants results in a higher contribution to the host country. This factor, alongside acculturation, impacts the attitudes of the native population towards immigrants, which in turn impacts immigration and naturalization policies. Poor integration of immigrants, on the other hand, may result in the

exclusion of immigrants and their children from the social and economic norms of the destination countries.

In the first chapter, I study whether the differential treatment towards rural-urban migrants in urban labor markets in China is limited only to that imposed by governmental policies based on the type of *hukou* registration, or whether this kind of treatment would still exist even after abolishment of the *hukou* system. This chapter offers a possibility to evaluate the actual impact of registration policies on the labor market integration of rural migrants in urban China by offering a more thorough *hukou*-based differentiation of individuals.

In the second chapter, I focus on two European countries: France and Denmark, and compare the returns to the personal characteristics of both naturalized and non-naturalized immigrants with those of the native population in the labor markets of the two destination countries. The analysis in this chapter not only allows us to determine the impact of naturalization on the labor market integration of immigrants, but it also attempts to estimate the impact of the strictness of naturalization policies on immigrant integration.

Finally, in the third chapter, I deviate from solely studying the benefits of naturalization and focus on the level of immigrant integration outside of the labor market context. In this chapter, I distinguish between the "social integration" of immigrants and the "social acceptance" of immigrants by the native population and link them to the public opinion of the native population about immigrants and immigration as well as the institutionally imposed restrictions on immigrants in 17 European countries.

Chapter 1: Differential Treatment in the Chinese Labor Market. Is *Hukou* Type the Only Problem?

1.1 Introduction and Literature Review

Differential treatment faced by minority groups in labor markets is one of the most important issues in labor economics, with immigrant groups (both internal and external) in destination labor markets being one of the minorities most studied by empirical researchers (Vink, Prokic-Breuer and Dronkers, 2013; Meng and Zhang, 2001). The importance of such research is in its potential application to citizenship and naturalization policy evaluations, a topic of broad and current interest in the era of globalization.

In China, differential treatment of minorities became important after 1978 when economic reforms were launched and internal migration, which used to be strictly prohibited, was allowed. According to the National Bureau of Statistics of China, the total number of rural-urban migrant workers reached around 200 million in 2008 and this number continues to grow. However, despite the reforms of the *hukou* system (household or residential regulation) in the 1980s and 1990s, migrant workers from rural areas in China still experience difficulty in acquiring local urban *hukou* registration and are often treated disadvantageously in urban labor markets by employers and governmental policies (Lu and Song, 2006).

Several studies have investigated the segmentation of the Chinese urban labor market based on the type of *hukou* registration, gender, occupational differences, type of business ownership, and income quintiles in the aggregate population (Meng and Zhang, 2001; Lee, 2012; Zhang and Wu, 2012). According to the related international literature, the findings of most empirical studies suggest that the Chinese labor market is also segregated based on the gender and race of workers (Magnani and Zhu, 2012). Moreover, by applying standard wage decomposition techniques and slightly adjusted versions of Mincer's (1958) wage equation on various data samples from China, most empirical studies in the field have also shown the existence of differential treatment towards rural *hukou* holders in urban labor markets (Song, 2013; Lee, 2012; Meng and Zhang, 2001; Zhang and Wu, 2012). However, the magnitude of such differential treatment varies highly based on the

chosen outcome variables, data sources, and market sectors. For example, using data from the population mini-census of China in 2005 and employing hourly earnings as the outcome variable, Zhang and Wu (2012) show that around 20% of the earnings disadvantages of rural *hukou* holders are attributable to occupational segregation. On the other hand, using data from five large cities in China, Lee (2012) finds that when adding bonuses and insurance contributions to the outcome variables, the difference in earnings unexplained by personal characteristics increases from 10% to 28%. Both of these empirical studies and most related empirical literature suggest the abolishment of the *hukou* system as a solution to eliminating of differential treatment. However, there is still a gap in the literature, which this chapter intends to fill.

In particular, it has not yet been empirically tested whether the differential treatment towards rural-urban migrants in the urban labor markets is limited only to that imposed by governmental policies based on the type of *hukou* registration, or whether this kind of treatment would still exist even after abolishment of the *hukou* system. According to the results of a survey conducted by the Chinese Academy of Social Science, nearly one third of Shanghai's population shows personal intolerance against migrants, indicating that they would not like to live next door to a migrant (The Economist, 2014). Such evidence may suggest that differential treatment towards rural-urban migrants in China is not entirely the result of the *hukou* registration system. In this chapter, I evaluate the actual impact of registration policies on the labor market integration of rural migrants in urban China by providing a more thorough *hukou*-based differentiation of individuals. The results of the evaluation suggest that the disadvantageous treatment towards rural migrants in urban labor markets in China is not entirely the result of the *hukou* registration system.

In respect of general migration theory, the possible implications of the results of this study may go beyond China. To my knowledge, this chapter presents one of the first attempts at differentiation between the negative attitude towards migrants imposed by governmental policies through the residential registration system, and the personal negative attitude of local employers and the local population. This issue is directly related to the immigrant integration literature, which has become popular in recent decades, particularly in case studies of European countries (Vink, Prokic-Breuer and Dronkers, 2013; Steinhardt, 2012). However, because of the difficulties associated with the acquisition of data regarding the non-registered migrants, the obtained estimates may often be biased. In the case of China, this problem may be solved with the use of

Chinese datasets and the *hukou* registration system, which is similar to the citizenship registration systems in many European and western countries.

1.2 The *Hukou* System

With the creation of a formal *hukou* system in China in 1958, the whole population was segmented into either rural or urban *hukou* holders, and there was a strict physical separation between the two groups. However, after the economic reforms of the 1980's and 1990's, the physical separation was eliminated and people were free to move between areas and change the category of their *hukou* (Song, 2014). Nevertheless, rural *hukou* holders in urban areas still have difficulties in attaining urban *hukou* and are often unfairly treated by employers and governmental policies in urban labor markets (Meng and Zhang, 2001).

According to Song (2014), based on the type of *hukou* registration in urban areas in China, all workers may be segmented into three major groups: those with permanent urban *hukou* registration; those with permanent rural *hukou* registration; and those who changed their *hukou* registration type. The latter can be further divided into those who changed their registration from urban to rural or rural to urban. However, despite the theoretical possibility of the former, it is seldom observed in practice because of the many advantages granted to urban *hukou* holders in the cities. Moreover, the segmentation also continues based on the place of registration (local vs. non-local) meaning that a local urban *hukou* holder in one city may also face some governmental restrictions in the labor market of another city (Song, 2014).

Local urban *hukou* holders in Chinese cities are granted advantageous social benefits by the city governments compared to rural or non-local *hukou* holders. These social benefits may, in general, be summarized as differences in access to certain employment positions, education, public retirement benefits, as well as medical and unemployment insurance (Démurger and Xu, 2013). However, these benefits are frequently changed and they vary from city to city, which makes it impossible to obtain a country representative measure of the differences in benefits. Unfortunately, because of the absence of freely available and publicly open information on the exact set of benefits and requirements for local urban *hukou*, it is not possible to specify these variations in order to

examine the importance of each separate factor for the successful integration of rural migrants. However, this issue is interesting and is an avenue for future research.

The advantageous treatment towards the local urban *hukou* holders tempts rural migrants to convert the type of their *hukou* from rural to local urban. However, despite the numerous refinements in the household registration system in China, there are still very few possibilities for rural migrants to obtain urban *hukou*. There are obviously also some costs associated with the loss of rural *hukou* status, i.e. the loss of rights to agricultural land and access to the social system in rural areas. Nevertheless, because of the presumable negligible impact of these costs in the context of urban labor markets, they are ignored in this study in accordance with the suggestions and methodology of related empirical literature (Zhang and Treiman, 2013; Démurger and Xu, 2013).

Quheng and Gustafsson (2014) summarize the three general ways of urban *hukou* attainment for rural migrants, which, however, are not fully sufficient for the *hukou* conversion and do not guarantee attainment of an urban *hukou*. These ways are: 1) Career routes through education, promotion, or joining the People's Liberalization Army or the Chinese Communist Party; 2) the so called "collective conversion", in which the agricultural land that belongs to individuals converts to urban use; 3) through joining the family of a local urban *hukou* holder, which is mostly applicable for women through marriage (Wong and Wai-Po, 1998). Several more recent routes that are effective in a number of cities cities are through a purchase of a house or even the direct purchase of *hukou*. However, these routes are seldom used and the information regarding the full set of requirements is not freely available (Zhang and Treiman, 2013).

"Collective converters" may be argued to be the most random and plausibly exogenous sample of the population compared to other converters. Moreover, "collective conversion" is the most commonly used route for *hukou* conversion, and even in my data sample, around half of the *hukou* converters reported using this route. Table 1.1a presents the average observable characteristics of *hukou* converters of working age in the data sample who changed their *hukou* status after 1990 by way of conversion. Unfortunately, information regarding the 3rd route is missing from the questionnaire.

Table 1.1a: Descriptive Statistics of Average Group Characteristics by the Route of *Hukou* Conversion.

Route of *Hukou* Conversion	Group Characteristics				Employment Status				N (as % of total converters)
	Age	Years of Education	Healthy	Female	Employed	Un-employed	Student	Home-maker	
Education/Promotion/ Army/Communist Party	30.93	13.50	86.97%	40.17%	88.88%	1.28%	8.13%	1.71%	467 (28.56%)
"Collective Conversion"	39.60	9.22	82.57%	50.33%	85.44%	4.55%	4.04%	5.97%	769 (46.92%)
Other	37.85	10.26	77.11%	55.97%	80.10%	7.21%	0.50%	12.19%	402 (24.52%)

In addition to being the most randomly chosen and plausibly exogenous group of *hukou* converters, the average observable characteristics of collective converters are also mostly in between those of the permanent urban population and migrants with rural *hukou*, indicating the high suitability of this group for the analyses (see tables 1.1b and 1.1c). Moreover, living on the borders of cities, these individuals represent a group of non-locals that is naturally integrated into the urban areas (since most probably they continued living their ordinary lives after the conversion of their *hukou* registration), which is also of benefit to the analyses.

There are very few empirical studies of differential treatment that consider more than two possibilities of *hukou* holders. For example, Cheng, Guo, Hugo and Yuan (2013) compared employment attainments and wage differentials between rural migrants, urban migrants and urban locals, and concluded that compared to urban locals, rural migrants suffer both employment and wage discrimination, whereas urban migrants suffer only wage discrimination. A similar study was conducted by Gagnon, Xenogiani and Xing (2014). These authors also concluded that both rural migrants and urban migrants are treated disadvantageously compared to urban locals, but this result was significant only outside of the formal labor market. Both of these findings contribute to the idea that the type of registration, even though it is important, does not guarantee elimination of disadvantageous treatment towards migrants in China. However, the results of these papers regarding the impact of the type of registration may not be very robust. This is because, although urban migrants are urban *hukou* holders, they are still different from urban natives and *hukou* converters by the place of their registration (local vs. non-local), which also imposes some governmental restrictions and policy-based differential treatment (Song, 2014).

Quheng and Gustafsson (2014) conducted perhaps the first empirical study of the characteristics that influence *hukou* conversion. They analyzed the economic well-being of the *hukou* converters by comparing them to those left behind in rural areas, and migrants who have kept their rural *hukou*. The authors concluded that there is a large incentive for the *hukou* conversion due to the high share of the differences in their incomes not related to their productivity. However, the authors did not conduct decomposition analyses to compare the urban locals to the *hukou* converters. The main difference between my study and the previous literature is that I test for the existence of differential treatment towards those rural migrants that converted their *hukou* type to local-urban, and thereby fully eliminated the disadvantageous treatment caused by governmental policies. I also compare the magnitude of this treatment to the one imposed on other rural migrants, thus obtaining an estimate of the pure "value" of residential registration.

1.3 Methodology

The theoretical model employed in the analysis is based on a more general Oaxaca and Ransom (1994) approach to the wage decomposition methodology of Blinder (1973) and Oaxaca (1973), which is heavily employed in economic literature to measure differential treatment in labor markets (Christofides, Polycarpou and Vrachimis, 2013; Song, 2013). The general procedure of the application of this model can be summarized in 3 steps. Firstly, the hourly earnings structures are separately estimated for individuals belonging to each of the three groups based on *hukou* registration type. In the second step, I estimate the "fair" hourly earnings structure of individuals, which is often referred to as the non-discriminatory income structure in the literature. Finally, I conduct decomposition analyses in order to measure the "unexplainable" component of the differences in hourly earnings.

The empirical model used to estimate the hourly earnings structures is similar to those widely employed in earnings decomposition analyses (Meng and Zhang, 2001; Song, 2013), which

are mainly based on Mincer's (1958) general capital earnings equation which may be specified as[1]:

$$\ln(Y_j) = \beta_0^j + \beta_1^j E + \beta_{12}^j E^2 + \beta_2^j TE + \beta_{22}^j TE^2 + \beta_3^j T + \beta_4^j O + \beta_5^j C + \beta_{6k}^j Ot\square er_k + \beta_{7l}^j PI_l + e_j, \quad (1.1)$$

where j=u, c, or r represents the sample groups of permanent urban *hukou* holders, *hukou* converters, and permanent rural *hukou* holders respectively; Y is the total hourly earnings (including bonuses); E is the years of education; TE is the current job tenure; T is a binomial dummy variable for training for the current job; O is the type of occupation (principal, technician, clerk); C is the type of employment contract (permanent, long-term); Other is a vector of other individual personal characteristics (k) including the school performance, marital status, height, body mass index, health condition, and ethnicity; PI is a set of binomial dummy variables for provinces and industries; and e is the error term.

One concern regarding this regression model may be the possibility that labor force participation for the majority or minority groups would be non-random. This would mean that there is a selection bias since I observe only the wages of those people who are working in the urban labor market and holding a certain type of *hukou* registration. Three types of selection bias may arise in the proposed analyses: (1) selection in *hukou* conversion among rural migrants, since they represent a sub-sample of migrants who chose to and succeeded in converting their *hukou;* (2) pre-migration selection, since the migrants represent a sub-sample of the rural population that chose to migrate; (3) selection into employment, since employed individuals represent a sub-sample of the potential labor force who chose to work. Selection bias may cause inconsistency in the results since the expected value of the error term in equation (1.1) may not be zero, given the fact of sub-sample selection.

The first type of selection bias is corrected through the use of only those *hukou* converters who obtained an urban *hukou* as a result of conversion of their agricultural land to urban use ("collective converters"). It is visible from the data (Table 1.1a) that *hukou* converters who use other conversion routes are on average younger individuals with a higher level of education and

[1] Multicollinearity of the variables was tested and ruled out from the model. The results of the tests are available upon request.

with selection patterns in terms of gender and employment status. Whereas "collective converters" represent a plausibly exogenous group, with average observable characteristics close to that of rural migrants. However, I cannot exclude the possibility that by initially living closer to cities the group of "collective converters" may also be closer to urban citizens in their observable characteristics compared to other individuals who initially had rural *hukou*.

In order to correct for the second type of selection I use the two-stage Heckman procedure for estimation of the income structure of migrants with rural *hukou*, incorporating those individuals who are left behind in rural areas. In the first stage, latent variables associated with the migration decision are used in order to estimate the associated parameter vectors, for which, in the second stage, the inverse mills ratios are then computed and included in the equation (1.1) for the group of rural *hukou* holders. In accordance with the related theoretical literature, I then take the estimated mills-term to the left-hand side of the regression equation and obtain the selectivity corrected income structure of migrants with rural *hukou*. Failure to control for migration selectivity, which has been often found by empirical literature (Xing, 2010), will result in an "unfair" comparison of rural migrants with the urban *hukou* holders, since the observable and unobservable characteristics of migrants that contribute to their migration decision may influence their incomes and result in overestimated returns to personal characteristics. The most important decision associated with the Heckman correction is the choice of the exclusion restriction variables in the first stage. In our case, these are the variables that impact the migration decision but do not impact the main outcome variable (i.e. the hourly earnings).

The exclusion restriction variables that were used in the first stage of the Heckman selectivity correction are the number of adult children (older than 20 years), and the number of elder siblings. The presence of children at the time of migration is generally considered to be an obstacle for migration, and I used the number of adult children at the time of the survey as a proxy variable. Even though the current number of children may arguably affect the current earnings through some channels, my logic was to use a proxy for the number of children at the time of migration rather than the current number of children. Clearly, the use of this proxy eliminates or at least reduces the number of possible channels of impact on the current earnings. This variable was additionally tested to have no direct effect on the hourly earnings, and proved to be a good exclusion restriction variable with a strong and significant negative impact on the migration

decision. On the other hand, the presence of elder siblings, which was my second exclusion restriction variable, is assumed to have a positive effect on the probability of migration by eliminating the obstacle of leaving the parents alone in the rural areas. It is worth mentioning that this sort of behavior of the elder children is culturally very specific to China, and our analysis also showed that the presence of elder siblings does have a statistically significant positive impact on the migration decision (Table 1.2a). Again, despite its possible impact on the hourly earnings through some channels (for example less investment in the human capital of each individual child by their parents), this variable was also tested to have no direct significant effect on the hourly incomes after controlling for other observable variables.

Despite all the presented arguments in favor of the chosen exclusion restriction variables, this still does not completely rule out the possibility of selectivity bias related to the migration decision in the estimations. There may still be some unobservable and uncontrollable (at least by our data) individual characteristics (e.g. the abilities of individuals, their ambitions, and character) that may hinder the migration decision and influence their incomes once employed in the urban labor markets. However, it is more reasonable to assume that the direction of this bias (if it exists and is not controlled for) will be more in favor rather against their competitiveness in the urban labor markets. This means that even in the presence of such bias, the results may underestimate of the value of the proper *hukou* registration status and the differences in the wage gaps between rural *hukou* holders and other individual groups.

Finally, I address the third selection issue and describe the offered solution later in the data description section.

The decomposition analyses cannot be operational without some assumptions about the structure of a "fair" (nondiscriminatory) estimate (β^*). The fair income structure is generally assumed to be something in between the structures of advantaged and disadvantaged groups and it may be represented by the following weighted equation:

$$\beta^* = \theta\hat{\beta}_u + \delta\hat{\beta}_c + (1-\theta-\delta)\hat{\beta}_r \text{ , with } \quad \theta, \delta \geq 0, \text{and } \theta + \delta \leq 1, \qquad (1.2)$$

where θ *and* δ are the weights, and any assumption about β^* may be reduced to an assumption about θ *and* δ. The theory proposes four general structures or values of the weights that may

clearly affect the resulting magnitude of differentials. Oaxaca (1973) proposes the adoption of either majority income structure or minority income structure as the fair estimates. On the other hand, Cotton (1988) proposes weights adjusted for the fractions of the groups included in the sample, reasoning that the fair structure should be more similar but not equal to that of the majority group. Finally, Neumark (1988) suggests that the fair estimate should be the one obtained from the pooled sample of all groups. The choice of the "fair" estimates does not impact the size of the wage gap, but may impact only the "unexplained" component of the gap. The analyses showed that this component was present under different choices of the "fair" estimates, but only the results based on the pooled sample estimate are reported[2].

Since the outcome variable includes not only the wages, but also the assigned bonuses, the unexplained component is not only a result of negative treatment towards the disadvantaged group, but also may be a result of the positive treatment towards the advantaged group. Thus, the pooled sample estimation of the "fair" component, which is also favored by the theoretical literature[3], was considered the most appropriate for the study[4].

Finally, the logarithmic decompositions of the (selectivity corrected) gross hourly earnings differentials at the means have the following forms:

$$\ln\overline{Y}_u - \ln\overline{Y}_r = (\overline{X}_u - \overline{X}_r)'\beta^* + \overline{X}_u'(\hat{\beta}_u - \beta^*) + \overline{X}_r'(\beta^* - \hat{\beta}_r), \qquad (1.3)$$

$$\ln\overline{Y}_u - \ln\overline{Y}_c = (\overline{X}_u - \overline{X}_c)'\beta^* + \overline{X}_u'(\hat{\beta}_u - \beta^*) + \overline{X}_c'(\beta^* - \hat{\beta}_c), \qquad (1.4)$$

$$\ln\overline{Y}_c - \ln\overline{Y}_r = (\overline{X}_c - \overline{X}_r)'\beta^* + \overline{X}_c'(\hat{\beta}_c - \beta^*) + \overline{X}_r'(\beta^* - \hat{\beta}_r), \qquad (1.5)$$

where $\overline{Y}_u, \overline{Y}_r$ and $\overline{Y}_c$ are the hourly mean total earnings of urban, rural, and converter workers respectively; β^* is the "fair" hourly earnings structure estimated by equation (1.2); $\overline{X}_u', \overline{X}_r'$ and $\overline{X}_c'$

[2] A common flaw of all the mentioned methods is that none of them accounts for possible changes in the wage structures under General Equilibrium as a result of changes in the supply of and demand for one type of workers in the labor markets.

[3] See Neumark (1988) and Oaxaca and Ransom (1994) for better descriptions of the pooled identification model and its advantages.

[4] Decomposition analyses based on other estimates of "fair" income structure provide similar results. These analyses are available upon request.

are the vectors of the mean values of the overall regressors from equation (1.1) respectively for urban, rural and converter workers; and $\hat{\beta}_u, \hat{\beta}_r$ and $\hat{\beta}_c$ are the conforming vectors of coefficients estimated by equation (1.1) for the corresponding data samples.

The first terms in the right-hand side of equations (1.3), (1.4) and (1.5) present estimates of productivity differentials or the income gap caused by different individual characteristics of the corresponding groups. The second terms estimate the unexplained advantage of the advantaged groups (permanent urban *hukou* holders in equations 3 and 4, and *hukou* converters in equation 5) due to different returns to personal characteristics of the corresponding groups. Finally, the third right-hand side terms in equations (1.3), (1.4) and (1.5) estimate the unexplained disadvantages of the minority groups which, together with the second terms, are generally considered to show the discrimination in the labor markets.

1.4 Data Description

For the empirical analyses I used one wave (2008) of the Longitudinal Survey on Rural Urban Migration in China (RUMiC[5]) data, which consists of three parts: the urban, rural, and migrant household surveys. I considered 2008 more appropriate in order for the results to be more comparable with other studies on RUMiC data that mostly used this wave, and in order to avoid the effect of a small positive shock in the level of unemployment which China experienced in 2009. A migrant is defined in these surveys as an individual who has a rural *hukou*, but who is living in an urban area at the time of the survey. RUMiC was designed to serve as a dataset for research that aims to observe migration patterns in China, covering the period 2008-2013 (only waves 2008 and 2009 are available) with surveys conducted annually. It surveys and includes

[5] The Longitudinal Survey on Rural Urban Migration in China (RUMiC) consists of three parts: the Urban Household Survey, the Rural Household Survey and the Migrant Household Survey. It was initiated by a group of researchers at the Australian National University, the University of Queensland and the Beijing Normal University, and was supported by the Institute for the Study of Labor (IZA), which provides the Scientific Use Files. The financial support for RUMiC was obtained from the Australian Research Council, the Australian Agency for International Development (AusAID), the Ford Foundation, IZA, and the Chinese Foundation of Social Sciences.

information concerning demographic characteristics, occupations, incomes, physical and mental health, and education from three groups of households: 5,007 rural–urban migrants who worked in 15 designated cities (Shanghai, Guangzhou, Shenzhen, Dongguan, Nanjing, Wuxi, Hangzhou, Ningbo, Wuhan, Chongqing, Chengdu, Hefei, Bangbu, Luoyang, and Zhengzhou) in 2008; 5,000 urban households in the same cities; and 8,000 rural households from 10 provinces or metropolitan areas where the 15 cities were located. Compared to other Chinese data sources, RUMiC has an advantage in containing migrant-representative data with fair migrant-rural and migrant-urban household proportions.

The main advantage of the data compared to other data sources is that, instead of conducting random interviews with migrants in selected urban neighborhood communities, the RUMiC research team employs a unique and effective sampling strategy to address the high concentration of rural-urban migrants in dormitories and construction sites surrounding rural suburbs[6].

For the purposes of this study, the data sample was restricted to include only those individuals of appropriate working age (16-60), excluding retired and unemployed individuals. Separate regression and decomposition analyses were performed for the sub-samples of males and females, as well as self-employed individuals[7]. The observations from the rural household survey were used only in order to correct for the migration selectivity of migrants. Finally, the whole sample of the urban work-force was divided into three sub-samples based on the *hukou* type as described in the methodology section.

Table 1.1b presents some descriptive statistics regarding the composition of the data by the type of *hukou* and gender. The table also addresses the concern of selection into employment described in the Methodology section. This study concerns only the individuals who are already employed in the labor markets, and selection into employment may be a concern only if there are significant group differences in them (i.e. differences in employment rates). Table 1.1b shows that

[6] More information can be found in Akgüç et al. (2014). "The RUMiC longitudinal survey: fostering research on labor markets in China". IZA Journal of Labor & Development, 3:5.

[7] The gender base separation and separate analysis of self-employed individuals are conducted because of the possibility differences in the wage structure between the groups (Hamilton, 2000; Giulietti, Ning and Zimmermann, 2012).

the group of rural *hukou* holders is the only "outlier" with employment rates being higher than those of the other two groups, which indicates lower employment selectivity. For the other two *hukou* possibilities, there is only a slight (1%) difference in the employment rates for males, and a 6.52% difference for females. These figures are consistent with the findings of related empirical literature of higher than 80% employment rates of all origin dependent groups in Chinese urban labor markets (Park and Wang, 2010), and speak in favor of a statement that this selection issue should not be problematic. Other studies using standard Heckman selectivity correction also suggest that labor selectivity has no significant effect on earnings (Lee, 2012).

Moreover, it is reasonable to assume that any selection into employment is based on some advantages in human capital. I assume that selection into migration (for which I control) will offset the phenomenon of higher employment rates among rural *hukou* holders, because the characteristics impacting the migration decision are roughly the same as those impacting the employment decision. Thus, selection into employment becomes even less of a concern after controlling for the migration decision.

Table 1.1b: Employment Status of Individuals of Working Age by Gender and Type of *Hukou*.

Gender	*Hukou* Type	Employed	Un-employed	Student	Home-maker	N
Male	Urban *Hukou*	92.72%	6.84%	0.06%	0.35%	3450
	Rural *Hukou*	99.06%	0.49%	0.32%	0.12%	4055
	Converted *Hukou*	93.72%	3.40%	1.05%	1.83%	382
Female	Urban *Hukou*	83.52%	9.58%	0.20%	6.67%	2955
	Rural *Hukou*	94.21%	1.56%	0.20%	3.86%	2954
	Converted *Hukou*	77.00%	5.68%	1.03%	16.28%	387

The main outcome variable in the regression analyses is the natural logarithm of hourly total earnings, which is computed by dividing the sum of total monthly wages and bonuses of individuals to the total amount of monthly hours worked. Unfortunately, when limiting the data sample to employed individuals who have no missing values of all variables used in the regression models, the number of observations decreased significantly, but it was still enough to obtain robust results.

The average values of the main outcome variable and the main set of explanatory variables of the employed urban population, which are used in the regression analyses, are presented in Table 1.1c. This set includes the length of tenure in the current job (individually reported years of work in current job), school performance (individually reported 1-5 scale scores of performance in their class before leaving school, with 1 standing for very good and 5 standing for very poor) and the health condition of individuals (individually reported 5-scale values with 1 standing for very good and 5 standing for very poor).

As expected, the individuals with the highest average total hourly earnings are those with permanent urban *hukou*, while the migrants with rural *hukou* experience on average the lowest hourly earnings among the three groups.

Table 1.1c: Descriptive Statistics of Average Individual Characteristics of Workers in the Paid-employment Sector by Gender and Self –employed Individuals by the Type of *Hukou*.

Employment and Gender		*Hukou* Type	Hourly Earning	Years of Education	School Performance	Age	Health	Current Tenure	N
Urban Paid Employment	Male	Urban *Hukou*	14.05	12.4	2.33	41.35	2.06	14.4	2063
		Converted *Hukou*	8.86	9.94	2.47	38.98	1.64	8.65	217
		Rural *Hukou*	7.21	9.44	2.85	31.23	1.71	3.66	2171
	Female	Urban *Hukou*	11.47	12.49	2.29	38.29	2.09	11.28	1744
		Converted *Hukou*	7.37	9.39	2.49	37.47	1.81	7.62	188
		Rural *Hukou*	6.09	9.15	2.77	30.11	1.81	2.81	1274
Urban Self-Employed	Male and Female	Urban *Hukou*	12.71	10.65	2.54	39.73	2	9.28	278
		Converted *Hukou*	11.27	9.86	2.57	37.82	2.11	6.46	27
		Rural *Hukou*	7.47	8.35	2.87	35.11	1.76	5.67	392
Rural	Male	Rural *Hukou*	7.79	8.54	2.69	35.25	1.79	7.81	4417
	Female	Rural *Hukou*	6.26	8.28	2.72	31.92	1.81	5.99	2490

Some group-dependent differences in the average characteristics are noticeable, particularly between the group of rural *hukou* holders and other groups. In particular, the rural *hukou* holders are on average younger and healthier individuals with lower levels of education and current job tenure compared to urban *hukou* holders. As I mentioned earlier, these differences are likely due to the selection to migration rather than selection to employment in urban labor markets. Moreover, when compared to the rural population, the migrants with rural *hukou* proved to be better selected for the employment market than those left behind in rural areas.

1.5 Empirical Results

In the first step of the analyses, the wage equations were estimated for migrants with rural *hukou* (Rural), migrants with converted *hukou* type (Converted), and individuals with permanent urban *hukou* (Urban), separately for males and females as well as self-employed individuals.

I started by addressing the issue of migration selectivity for migrants without urban *hukou* registration, the results of which are reported in Table 1.2a for individuals in the paid-employment sector, and in the 2nd and 3rd columns of Table 1.2b for self-employed individuals. The dependent variables in all earnings equation models are the natural logarithms of total hourly earnings (including bonuses).

The positive sign and significance of the coefficients of inverse mills ratios show that, first, the observable and unobservable characteristics of individuals that induce the migration decisions of the rural population positively influence their incomes in urban labor markets. Secondly, I find that the control for migration selectivity is necessary for both the paid-employment sector and self-employed individuals.

The second step of the analyses was to estimate the income structures for the other two groups based on their *hukou* status, the results of which are reported in Table 1.2c for the paid-employment sector and in the 1st columns of Table 1.2b for self-employed individuals. The regression outputs show that the income structures of the observed groups are very different from one another, and the separation of the data based on gender and employment sector are also necessary to obtain unbiased results.

The coefficients of the important explanatory variables are mostly highly significant and have the expected values suggested by other empirical literature. The values of R squared are reasonable, indicating that the obtained income structures are sufficient for use in income decomposition. The main differences between the total income determination for urban residents and rural migrants with and without urban *hukou* registration are as follows. First, when analyzing the paid-employment sector, I notice that the constant terms are measurably lower for the rural migrants without urban *hukou*, indicating that there may be disadvantageous differential treatment towards them. Secondly, the return to current job tenure is higher for rural

Table 1.2a. Heckman Correction 1st and 2nd Stage Regression Results for Migrants with Rural *Hukou* by Gender.

VARIABLES	(1) Rural (M)	(2) 1st Stage (R-M)	(3) Rural (F)	(4) 1st Stage (R-F)
Years of Education	-0.013	-0.022	0.057***	0.039
	(0.020)	(0.045)	(0.022)	(0.042)
Years of Education Squared	0.003***	0.006***	-0.000	0.004*
	(0.001)	(0.002)	(0.001)	(0.002)
Ethnic Minority	-0.030	-0.242	-0.132	-0.631***
	(0.062)	(0.163)	(0.095)	(0.164)
School Performance	0.020	0.272***	-0.015	0.132***
	(0.016)	(0.037)	(0.021)	(0.040)
Married	-0.040	-0.286***	-0.180***	-0.366***
	(0.028)	(0.072)	(0.039)	(0.081)
Health	-0.037***	-0.132***	-0.016	-0.094***
	(0.014)	(0.034)	(0.018)	(0.035)
Body Mass Index (bmi)	0.010	-0.065***	-0.015*	-0.061***
	(0.006)	(0.015)	(0.008)	(0.016)
Height	1.385***	5.100***	1.065***	0.222
	(0.226)	(0.454)	(0.273)	(0.509)
Age	0.041***	-0.005	0.051***	0.094***
	(0.008)	(0.019)	(0.012)	(0.021)
Age Squared	-0.001***	0.000	-0.001***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Recent Training	0.055***		0.056*	
	(0.021)		(0.029)	
Current Tenure	0.040***		0.040***	
	(0.005)		(0.009)	
Current Tenure Squared	-0.001***		-0.001	
	(0.000)		(0.001)	
Number of Adult Children		-0.695***		-0.432***
		(0.083)		(0.074)
Number of Elder Siblings		0.091***		0.072***
		(0.018)		(0.016)
Mills Term	0.670***		1.006***	
	(0.154)		(0.139)	
Constant	-1.779***	-7.327***	-1.225**	-0.826
	(0.445)	(0.925)	(0.534)	(1.004)
F-statistics for IV (q=2)		21.63***		5.97***
Observations	2,171	6,605	1,274	7,509

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The left-hand side variables in all final regression models are the total hourly earnings of the specified groups. The first stage regression models control for the migration decision of the rural *hukou* holders in urban labor markets. All regression models also include sets of binomial dummy variables including occupations, industries, job sectors, and provinces.

Table 1.2b. OLS and Heckman Correction 1st and 2nd Stage Regression Results for Self-employed Individuals by their Type of *Hukou* Registration.

	(1)	(2)	(3)
VARIABLES	Permanent Urban	Permanent Rural	1st Stage
Years of Education	0.0615**	0.022	-0.008
	(0.0311)	(0.06)	(0.049)
Years of Education Squared	-0.00139	0.000	0.000
	(0.00103)	(0.004)	(0.003)
Ethnic Minority	0.140	-0.516*	-0.462**
	(0.375)	(0.284)	(0.216)
School Performance	-0.0287	0.131**	0.116***
	(0.0609)	(0.057)	(0.045)
Married	0.385**	0.360***	0.306***
	(0.172)	(0.134)	(0.102)
Health	-0.0119	-0.184***	-0.182***
	(0.0587)	(0.052)	(0.040)
Body Mass Index (bmi)	0.0104	-0.036*	-0.052***
	(0.0234)	(0.022)	(0.017)
Height	1.592**	-2.742***	-5.753***
	(0.622)	(0.990)	(0.448)
Age	0.0583	0.078**	0.156***
	(0.0415)	(0.040)	(0.026)
Age Squared	-0.000875*	-0.001**	-0.002***
	(0.000506)	(0.001)	(0.000)
Recent Training	0.0874	0.073	
	(0.114)	(0.113)	
Current Tenure	0.0367**	0.025*	
	(0.0173)	(0.015)	
Current Tenure Squared	-0.000474	-0.001*	
	(0.000527)	(0.001*)	
Number of Adult Children			-0.325***
			(0.073)
Number of Elder Siblings			-0.009
			(0.017)
Mills Term		1.400***	
		(0.243)	
Constant	-2.804**	3.927***	6.671***
	(1.295)	(1.505)	(0.913)
F-statistics for IV (q=2)			11.65***
Observations	278	392	4825
R-squared	0.408		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The left-hand side variables in all final regression models are the total hourly earnings of the specified groups. The first stage regression models control for the migration decision of the rural *hukou* holders in urban labor markets. All regression models also include sets of binomial dummy variables including occupations, industries, job sectors, and provinces.

Table 1.2c. OLS Regression Results for Individuals with Permanent Urban *Hukou* and Rural Migrants with Converted *Hokou* Type by Gender.

VARIABLES	(1) Urban (M)	(2) Urban (F)	(3) Converted (M)	(4) Converted (F)
Years of Education	0.0649***	0.0796***	-0.0460	0.0714
	(0.0135)	(0.0146)	(0.0567)	(0.0601)
Years of Education Squared	-0.00114**	-0.00176***	0.00206	-0.00264
	(0.000459)	(0.000467)	(0.00263)	(0.00315)
Ethnic Minority	-0.0848	0.148	-1.112**	-0.488
	(0.109)	(0.126)	(0.492)	(0.693)
School Performance	-0.0433***	-0.0529***	-0.103*	0.0280
	(0.0165)	(0.0184)	(0.0583)	(0.0590)
Married	0.147***	0.0275	-0.102	-0.0249
	(0.0396)	(0.0366)	(0.147)	(0.148)
Health	-0.00932	0.00459	0.0900	0.0461
	(0.0169)	(0.0179)	(0.0547)	(0.0557)
Body Mass Index (bmi)	0.0190***	-0.0166**	0.0194	-0.00991
	(0.00695)	(0.00728)	(0.0207)	(0.0246)
Height	0.0155	0.563**	0.713	-0.305
	(0.237)	(0.270)	(0.709)	(0.769)
Age	0.00240	0.0133	0.00828	-0.00125
	(0.0117)	(0.0132)	(0.0359)	(0.0396)
Age Squared	-0.000106	-0.000213	-0.000227	-0.0000482
	(0.000140)	(0.000168)	(0.000432)	(0.000505)
Recent Training	0.0175	0.115***	-0.0354	0.202**
	(0.0238)	(0.0254)	(0.0770)	(0.0908)
Current Tenure	0.0200***	0.0213***	0.0115	-0.00780
	(0.00439)	(0.00444)	(0.0134)	(0.0188)
Current Tenure Squared	-0.000342***	-0.000371***	0.000144	0.000204
	(0.000116)	(0.000129)	(0.000396)	(0.000728)
Permanent Contract	0.470***	0.357***	0.233	0.0722
	(0.0402)	(0.0420)	(0.164)	(0.170)
Long-term Contract	0.339***	0.199***	0.135	-0.0662
	(0.0346)	(0.0335)	(0.111)	(0.103)
Constant	0.585	0.152	0.643	1.866
	(0.495)	(0.534)	(1.490)	(1.562)
Observations	2,063	1,744	217	188
R-squared	0.434	0.428	0.408	0.428

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The left-hand side variables in all final regression models are the total hourly earnings of the specified groups. All regression models also include sets of binomial dummy variables including occupations, industries, job sectors, and provinces.

migrants without urban *hukou* registration, which could be because the current job tenure is evaluated highly during the first years after the job attainment and rural migrants had very limited work experience in urban areas. Finally, the return to physical abilities (expressed by age, body mass index, height and health of individuals) are mostly significant and measurably higher for migrants with rural *hukou*, whereas the returns to mental abilities (expressed by years of education and school performance) are mostly significant and higher for urban *hukou* holders (both permanent and converters). This may indicate that migrants with rural *hukou* registration are mostly concentrated in the market sector where physical abilities are evaluated higher, whereas urban *hukou* holders are concentrated in the market sector, which evaluates mental abilities more highly.

For self-employed individuals, I obtain roughly similar results (Table 1.2b). Unfortunately, the small number of observations does not allow us to obtain income structures by gender as well as that of self-employed *hukou* converters, and conduct a three-level income decomposition based on the type of *hukou*. Therefore, I conduct only a two-level (Urban-Rural) income decomposition for the sample of self-employed individuals with no gender differentiation. However, since the advantages granted by *hukou* registration are mostly bound to the paid-employment sector, I expect that the possible disadvantageous treatment imposed solely by governmental restrictions based on the type of *hukou* would be very low for this sample, and I also do not expect much gender differentiation in this sector.

The next step in the decomposition analyses was to estimate the "fair" income structure or the income structure that would have existed in the absence of any sort of unexplainable disadvantageous treatment. The literature suggests four general possibilities for estimation of the "fair" income structure. However, I focus only on the joint income decomposition model, since I believe that it presents the most realistic results. The results of the joint regression models are in Table 1.4 in the Appendix.

Finally, the last step of the analyses is the hourly total earnings decomposition itself. Table 1.3 presents the results of the decomposition analyses for all 3 sub-samples, and with the separation based on the employment sector and gender in the paid-employment sector. The results show statistically and economically significant "unexplained" components in the income gaps between

all the observed groups except for the case of the Converted-Rural comparison for males working in the paid-employment sector. However, there are some specifications that should be mentioned in more detail.

First, the Urban-Rural differences in total hourly earnings for females employed in the paid-employment sector are higher than those for males, and they are almost equally divided between the Urban-Converted and Converted-Rural couples. For males working in the paid-employment sector, a higher part of the difference in Urban-Rural total earnings is captured by the Urban-Converted difference, which may indicate that the type of *hukou* registration has a more measurable impact on the labor market earnings of females rather than males.

Table 1.3. Total Earnings Decomposition Results.

EMPLOYMENT AND GENDER	Paid-employment (Male) Group 1 - Group 2			Paid-employment (Female) Group 1 - Group 2			Self-employed
	Urban - Rural	Urban - Converted	Converted - Rural	Urban - Rural	Urban - Converted	Converted - Rural	Urban - Rural
ESTIMATES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted Average Earnings of Group 1	2.442	2.442	2.036	2.242	2.242	1.85	2.281
	(0.014)	(0.014)	(0.036)	(0.015)	(0.015)	(0.038)	(0.045)
Predicted Average Earnings of Group 2	1.734	2.036	1.734	1.421	1.85	1.421	1.801
	(0.041)	(0.036)	(0.041)	(0.045)	(0.038)	(0.045)	(0.019)
Difference in Average Earnings	0.707***	0.406***	0.302***	0.821***	0.392***	0.429***	0.480***
	(0.044)	(0.039)	(0.055)	(0.048)	(0.041)	(0.059)	(0.049)
Explained Component of the Difference	0.584***	0.306***	0.278***	0.570***	0.302***	0.268***	0.390***
	(0.016)	(0.030)	(0.029)	(0.018)	(0.030)	(0.031)	(0.040)
Unexplained Component of the Difference	0.124***	0.0993***	0.024	0.251***	0.0895**	0.161***	0.0895***
	(0.041)	(0.033)	(0.051)	(0.044)	(0.036)	(0.056)	(0.029)
Observations	4,234	2,280	2,388	3,018	1,932	1,462	670

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Secondly, the results show that around 80% of the male and around 70% of the female earnings gaps between individuals holding urban and rural *hukous* come from differences in observable characteristics (explained component). When observing the unexplainable components of the differences between the three groups, I can see that for females working in the paid-employment sector a higher share of the unexplained component of the Urban-Rural difference is captured by the Converted-Rural group. However, roughly one-third of the unexplainable component in the

Urban-Rural earnings gap, which is around 10% of the entire Urban-Rural earnings gap, is concentrated in the Urban-Converted difference. This may indicate that the governmental restrictions based on the type of *hukou* are responsible for 2/3 of the unexplainable disadvantageous treatment in hourly earnings of females employed in paid-employment sector or around 20% of the entire Urban-Rural earnings gap. Nevertheless, it should be noted that the significant unexplainable component in the Converted-Rural earnings gap may also be a result of positive selection to employment of the group of converted *hukou* holders discussed in the methodology section (Table 1.1b).

The results for males working in the paid-employment sector show that there is a very small unexplained component in the Converted-Rural differences in hourly earnings, which is not even statistically significant. This indicates that the differential treatment against the rural migrants belonging to this group is not a result of the imposed governmental restrictions based on the type of *hukou* registration. In general, the results suggest that around 20% of the entire Urban-Rural earnings gap for males working in the paid-employment sector is rather a result of the personal attitude of urban employers towards non-locals.

This finding may also be supported by the results obtained from the sample of self-employed individuals, which also show around 20% of unexplainable gaps in earnings. Since the self-employed individuals are not restricted in the labor markets by the *hukou* regulation, it is reasonable to assume that the Urban-Rural difference for this sample does not capture the impact of governmental restrictions based on the type of *hukou* registration. The treatment toward this group of individuals in the Chinese labor market has not been sufficiently studied in previous literature, and the findings suggest that origin-dependent differential treatment exists even here.

In general, the findings of this chapter are in line with the findings of the previous literature, suggesting that the type of *hukou* registration of rural migrants does have an influence on their total hourly earnings, particularly for the females working in the paid-employment sector. However, the findings also suggest that the *hukou* registration system is not entirely responsible for the emergence of an unexplained component in the income differences, and even the abolishment of this system would at most be able to eliminate only one share of the differential treatment towards migrants.

1.6 Conclusion

I introduce a new dimension for income decomposition in China, which will add to the understanding of the possible effects of changes in governmental policies on differential treatment in urban labor markets in China. In particular, I test whether there is differential treatment towards those individuals in urban China who converted the type of their official registration (*hukou*) to local-urban registration, and how it differs from the treatment towards those who did not convert their registration type.

The findings indicate that in the given data sample, the type of *hukou* is not entirely responsible for the unexplainable disadvantageous treatment towards migrants working in the paid-employment sector, since measurable unexplained components in the total income differences persist for the Converted-Rural couples for workers of both genders (males and females). Moreover, the results suggest that unexplainable disadvantageous treatment against migrant males working in the paid-employment sector may even not decrease at all after their *hukou* registration is converted from rural to urban. One explanation for this phenomenon may be a possibility that, conditional on the conversion of their *hukou* status, this group tends to become employed in a sector of the urban market that experiences softer differential treatment. However, the construction of the data does not allow us to test this hypothesis.

Despite its considerable assumed impact on the differential treatment towards rural migrant workers in urban China, the type of *hukou* registration is not fully responsible for unexplainable disadvantageous treatment towards non-locals. This means that reforms in the *hukou* registration system or even the complete abolishment of this system, which has been one of the main recommendations of empirical researchers, will not actually cause the elimination of the differential treatment in Chinese labor markets.

The *hukou* system is generally considered to have many similarities to migrant registration systems of various European countries. These similarities open a wide range of prospects for research and generalization of the results both from Chinese to European markets and vice-versa. However, for generalization of the results of a complicated issue such as differential treatment, additional analyses need to be conducted also for European countries.

1.A Appendix 1

Table 1.4. OLS Regression Results for the Joint Samples with All 3 *Hukou* Possibilities.

VARIABLES	(1) Joint (Male)	(2) Joint (Female)	(3) Joint (Self-employed)
Years of Education	0.0447***	0.0630***	0.0380**
	(0.00844)	(0.00898)	(0.0161)
Years of Education Squared	-0.000372	-0.00118***	-7.02e-05
	(0.000330)	(0.000336)	(0.000692)
Ethnic Minority	-0.0299	0.0767	-0.00712
	(0.0530)	(0.0725)	(0.145)
School Performance	-0.0267**	-0.0486***	-0.00623
	(0.0104)	(0.0125)	(0.0263)
Married	0.0471**	-0.0378	0.0917
	(0.0224)	(0.0244)	(0.0633)
Health	-0.00756	-0.00369	-0.0195
	(0.0101)	(0.0114)	(0.0232)
Body Mass Index (bmi)	0.0178***	-0.00958*	0.00766
	(0.00438)	(0.00497)	(0.0102)
Height	0.519***	0.787***	1.131***
	(0.141)	(0.176)	(0.254)
Age	0.0255***	0.0262***	-0.0148
	(0.00550)	(0.00700)	(0.0158)
Age Squared	-0.000390***	-0.000395***	0.000124
	(7.02e-05)	(9.42e-05)	(0.000205)
Recent Training	0.0218	0.0804***	0.166***
	(0.0154)	(0.0178)	(0.0546)
Current Tenure	0.0256***	0.0239***	0.0175**
	(0.00271)	(0.00323)	(0.00753)
Current Tenure Squared	-0.000388***	-0.000417***	-0.000261
	(7.80e-05)	(0.000101)	(0.000295)
Permanent Contract	0.210***	0.215***	
	(0.0222)	(0.0260)	
Long-term Contract	0.169***	0.155***	
	(0.0172)	(0.0193)	
State Owned Enterprises	0.0482**	0.0370*	
	(0.0190)	(0.0221)	
Constant	-0.377	-0.301	-0.329
	(0.276)	(0.333)	(0.544)
Observations	4,660	3,660	1,408
R-squared	0.467	0.453	0.201

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The left-hand side variables in all final regression models are the total hourly earnings of the specified groups. All regression models also include sets of binomial dummy variables including occupations, industries, job sectors, and provinces.

Chapter 2: Treatment-Related Naturalization Premiums in Two European Countries: Evaluation and Comparison

2.1 Introduction and Literature Review

In recent decades, studies of the determinants, causes, and benefits of naturalization have created a well-developed body of research for many countries (Bratsberg, Ragan and Nasir, 2002; DeVoretz and Pivenko, 2004). Some of these studies focus on the impact of various individual characteristics, as well as origin and destination country characteristics, on the obtainment of citizenship in destination countries. Others attempt to estimate the socio-economic benefits of naturalization via testing for the impact of destination country citizenship on employment probabilities and employment wages.

In economic literature, citizenship is defined as a legal status that expresses the state-individual relationship, and grants some rights to those who hold it, which include, among others, physical and political protection and the right to vote. These rights clearly promote the popularity of naturalization, which is defined as citizenship acquisition for immigrants in the countries of destination. Moreover, the act of naturalization also serves as a signal for potential employers that the applicant intends to stay and work in the country of destination. This signal, along with the reduced administrative costs from the employer's perspective, may increase both the chances of employment for naturalized immigrants and their employment wages once employed (Vink, Prokic-Breuer and Dronkers, 2013; Steinhardt, 2012).

Many studies of the benefits of naturalization have been conducted for countries that serve as the most popular destinations for immigration, e.g. the US (Bratsberg, Ragan and Nasir, 2002), Canada (DeVoretz and Pivenko, 2004), and many European countries including Germany (Steinhardt, 2012), Sweden (Engdahl, 2011), France (Fougere and Safi, 2008) and the Netherlands (Bevelander and Veenman, 2008). Most of the literature shows strong and significant positive naturalization premiums in the observed countries, even after controlling for individual characteristics (Bratsberg, Ragan and Nasir, 2002; DeVoretz and Pivenko, 2004; Fougere and Safi, 2008). Another important common conclusion in most literature is that the naturalization

premiums are generally greater for males (Steinhardt, 2012), and for immigrants from less-developed countries who have lower employment probabilities (Bratsberg, Ragan and Nasir, 2002; Fougere and Safi, 2008).

In this chapter, I focus on two European countries: France and Denmark. The choice of the two countries was based on their similarities of the countries with regard to country specific characteristics (e.g. geographical location), and similar shares of EU-born and non EU-born immigrants. Moreover, the choice of the two countries was further supported due to the availability of the necessary variables in the data (see the "Data" section later), and the differences in their naturalization policies. Based on the requirements and conditions for naturalization, Denmark can be considered a country with strict naturalization policies, whereas France has relatively softer policies[8]. The differences in the naturalization policies are necessary for this study in order to make conclusions regarding the role of the strictness of naturalization policies in the results.

Studies of the determinants of naturalization mostly conclude that naturalization rates are positively influenced by the softness of citizenship policies in the countries of destination. In particular, the literature looks at two important characteristics of citizenship policies i.e., the tolerance towards dual citizenship and the minimal required duration of residency in the destination countries. There is some empirical evidence of the negative relationship between the minimal required duration of residency and the rates of naturalization, whereas other empirical evidence suggests a positive relationship between the tolerance towards dual citizenship and naturalization rates (Vink, Prokic-Breuer and Dronkers, 2013).

While there is sufficient evidence of a positive relationship between naturalization and economic outcomes, there are very few studies that conduct cross-county comparisons to determine the effect of macroeconomic factors and institutional policies on the level of immigrant integration. In addition, few studies focus on the institutional determinants of the immigrant-native gaps in labor market performance. For example, using longitudinal data on native-immigrant gaps in labor market performance in 21 European countries, Guzi, Kahanec and Kureková (2015) show that the types of Varieties of Capitalism (VoC) framework matter for immigrant integration in the

[8] A more detailed description of the citizenship policies of both countries is found in Appendix 2.

host labor markets. The authors find that in countries with mixed market economies immigrants perform better in terms of labor force participation and permanent employment. Moreover, in their study, Guzi and Kahanec (2017) also show that in many European countries certain groups of immigrants may sometimes be even more flexible than the natives in their response to labor market shocks.

In this study, however, I focus on the impact of the citizenship policies on the level of immigrant integration, which is also a topic not covered enough by cross-country empirical research. One example of a cross-country comparison study is that conducted by Bevelander and Pedakur (2012) for Canada and Sweden, in which the authors showed that naturalization has a positive impact on both the employment probabilities and relative incomes of immigrants, and that this impact is stronger in Sweden, which has very low barriers to naturalization. On the other hand, in their comparison of Denmark and Sweden Helgertz, Bevelander and Tegunimataka (2014) again considered Sweden the country with lower barriers to naturalization and obtained very similar naturalization premiums for both countries. A more comprehensive study of the relationship between citizenship policies and naturalization was conducted by Vink, Prokic-Breuer and Dronkers (2013) for 16 European countries. The authors showed that more accessible citizenship policies increase the naturalization rates, but this finding was significant only for immigrants from less developed countries.

To my knowledge, there is no empirical comparison of the returns to personal characteristics of both naturalized and non-naturalized immigrants with those of the native population in labor markets at destination countries. In this chapter, I conduct such a study to determine the impact of naturalization on labor market integration, rather than simply estimate the employment and wage premiums of naturalization. Moreover, the analyses are conducted for two European countries, France and Denmark, that have different policies for naturalization, which allows us to attempt to obtain an estimate of the impact of naturalization policies on the treatment in labor markets after naturalization.

2.2 Methodology

The theoretical model employed in the analyses is the more general Oaxaca and Ransom (1994) approach to the wage decomposition methodology by Blinder (1973) and Oaxaca (1973), which is a model heavily employed in the empirical literature to evaluate treatment towards workers in labor markets (Beblo, Beninger, Heinze and Laisney, 2003). This model has several advantages compared to a simple OLS model with citizenship status-related dummy variables. First, the Oaxaca-Ransom model allows the estimation of the impact of the specific group-belonging factor, not only through the effect on the constant variable but also through the direct or indirect influence of all other individual characteristics. In other words, instead of having only one dummy variable for group-belonging, this model is similar to having a whole set of dummy variables interacted with the entire set of explanatory, as well as constant variables. Secondly, this model not only shows the advantage or disadvantage of the group-belonging factor, but it is also able to decompose that gap into two parts compared to a chosen threshold (in this analyses, this threshold is the "fair" wage structure). By such decomposition, if the threshold is chosen properly, this model allows us to distinguish between the advantage of the advantaged group and the disadvantage of the disadvantaged group, providing more insights into the formation of the wage gaps and allowing more robust conclusions[9].

The data from each country (described in the next section) was divided into three sub-samples: "eligible" migrants (immigrants who are not naturalized but who reside in the countries for longer than the minimal required duration for naturalization in the country of residence); naturalized migrants (immigrants who are naturalized); and natives (citizens of the country of residence that were born in the country of residence). The data was further separated based on the gender of individuals and the analyses were conducted separately for each gender[10]. Through this differentiation, every individual was categorized into one of the 12 groups or sub-samples based on the country (France and Denmark), gender (male and female), and citizenship status (eligible migrant, naturalized migrant, and native).

[9] For more detailed description of the advantages of this model, see Beblo, Beninger, Heinze and Laisney (2003).

[10] The gender-based separation is conducted to control for different wage structures and occupational attainment probabilities of different genders.

The most appropriate procedure of the employed model is summarized in 3 steps. In the first step, the employment wage structures for individuals belonging to each of the three groups, based on the type of their citizenship and naturalization status (i.e. naturalized immigrants, non-naturalized immigrants, and native population), was separately estimated for each gender and country of interest. In the second step, a "country-equalized" wage structure (one wage structure for each gender and country) was constructed, which is the wage structure that would have been applicable for all three groups in a particular country in the absence of any sort of treatment or benefit differences. In the final steps, the decomposition analysis was conducted to measure the unexplainable component in the wage differences that is caused by unequal returns to personal characteristics.

In this analysis, for each country I use a version of Mincer's (1958) general capital earnings equation to estimate the wage structures for individuals belonging to each of the three groups based on the type of their citizenship and naturalization status. This identification model is heavily used in the empirical literature in earnings decomposition analyses and is specified by the following equation:

$$\ln(Y_{cj}) = \beta_0^{cj} + \beta_1^{cj} E + \beta_2^{cj} T + \beta_3^{cj} O + \beta_4^{cj} I + \beta_{5k}^{cj} Ot\square er_k + e_{cj}, \quad (2.1)$$

where c represents the country, j=c, n, or m represents the sample groups of the native population, naturalized immigrants, and non-naturalized immigrants, respectively; Y is the hourly employment wage of individuals; E is the years of education; T is the current job tenure (the age of individuals is used as a proxy); O is the type of occupation (managers, skilled workers, and laborers); I is other job information (size of the firm, or number of subordinates); Other is a vector of other individual personal characteristics (k) including marital status and the number of children; and e is the error term.

The main concern with this regression model is the possibility of selection bias, particularly for naturalized immigrants who may be a non-random sample of the immigrants. Such bias may arise if naturalized and non-naturalized migrants pass through some selection processes based on certain unobservable characteristics, which later influence their employment wages or employment probabilities. Unfortunately, I was not able to correct for such selection issues because of the limitations in the data, which give rise to some concerns regarding the interpretations of the results.

However, the analysis relies on the conditional independence assumption (CIA)[11], which amounts to the assumption that, conditional on the observable characteristics, the selection bias disappears. There is generally no technical solution to ensure that CIA holds or to disprove the applicability of CIA for the analysis. However, since naturalization (the granting of citizenship status) is based solely on the observable characteristics of individuals, it is reasonable to assume that the potential outcome is independent of the naturalization status once controlled for the observable characteristics. The selection issues, their possible impacts on the findings, and some explanations as to why I can neglect them are discussed in more detail later in the result section.

The decomposition analysis cannot become operational without some assumptions about the structure of an "equalized", or in this case "country-equalized", estimate of wage structure (β^*). This wage structure is generally assumed to be something in-between the structures of advantaged and disadvantaged groups, and is represented by the following weighted equation:

$$\beta_C^* = \theta\hat{\beta}_{Cc} + \delta\hat{\beta}_{Cn} + (1-\theta-\delta)\hat{\beta}_{Cm}, \quad \text{with} \quad \theta, \delta \geq 0, \text{and } \theta + \delta \leq 1, \tag{2.2}$$

where θ and δ are the weights, and any assumption about β_C^* may be reduced to an assumption about θ and δ. In the decomposition analysis, I use the method based on Neumark (1988), and Oaxaca and Ransom's (1994) suggestion that this estimate should be obtained from the pooled sample of all groups.

Finally, the logarithmic decompositions of the employment wage differentials at the means for each country have the following forms:

$$\ln\overline{Y}_c - \ln\overline{Y}_m = (\overline{X}_c - \overline{X}_m)'\beta^* + \overline{X}_c'(\hat{\beta}_c - \beta^*) + \overline{X}_m'(\beta^* - \hat{\beta}_m), \tag{2.3}$$

$$\ln\overline{Y}_c - \ln\overline{Y}_n = (\overline{X}_c - \overline{X}_n)'\beta^* + \overline{X}_c'(\hat{\beta}_c - \beta^*) + \overline{X}_n'(\beta^* - \hat{\beta}_n), \tag{2.4}$$

$$\ln\overline{Y}_n - \ln\overline{Y}_m = (\overline{X}_n - \overline{X}_m)'\beta^* + \overline{X}_n'(\hat{\beta}_n - \beta^*) + \overline{X}_m'(\beta^* - \hat{\beta}_m), \tag{2.5}$$

[11] See Rubin (1991): "Practical Implications of Modes of Statistical Inference for Causal Effects and the Critical Role of the Assignment Mechanism".

where $\overline{Y}_c$, $\overline{Y}_n$ and $\overline{Y}_m$ are the mean hourly employment wages of the native population, naturalized immigrants, and non-naturalized immigrants, respectively; β^* is the "country-equalized" employment wage structure estimated by equation (2.2) for each country; $\overline{X}'_c$, $\overline{X}'_n$ and $\overline{X}'_m$ are the vectors of the mean values of the overall regressors from equation (2.1) for the native population, naturalized immigrants, and non-naturalized immigrants respectively; and $\hat{\beta}_c$, $\hat{\beta}_n$ and $\hat{\beta}_m$ are the conforming vectors of coefficients estimated by equation (2.1) for the corresponding data samples.

The first terms in the right-hand side of equations (3), (4), and (5) present estimates of productivity differentials or the wage gap caused by different average individual characteristics between the corresponding groups. The second and third terms estimate the "unexplained" components of the difference caused by differentiated returns to individual characteristics. These components are called "unexplained" because they are not caused by the difference in individual characteristics but are rather caused by the difference in returns to these characteristics. The estimation of the values and the significance of these "unexplained" components for each of the country-based (also gender-based) sub-samples is the main aim of the decomposition analysis.

2.3 Data

The main results of the chapter were obtained using the 2010 data samples for France and Denmark from the Luxemburg Income Study (LIS) database. The LIS includes, among other things, personal-level micro-data on demographic characteristics, employment, and individual incomes from countries in Europe, North America, Latin America, Africa, Asia, and Australasia.

As I mentioned in the introduction, the choice of France and Denmark for the analysis, among other reasons, was based on the availability of the necessary characteristics in the data (the most important of which were the citizenship and the duration of residency in the country of destination for migrants), and the differences in the citizenship policies in the two countries. In this study, Denmark is considered a country with strict citizenship policies, whereas France has

softer policies[12]. The strictness of the citizenship policies was measured based on the tolerance towards dual citizenship, the minimal required duration of residency, and other requirements[13].

As described in the methodology section, the data from each country is divided into three sub-samples based on their origins and naturalization status (natives, naturalized, and eligible migrants), and the analyses are conducted separately for each gender. The data was also limited to include only those individuals of appropriate working-age (25 to 54 years old) and those first generation immigrants who are eligible for naturalization (i.e. who have lived at least the minimal required durations in the destination countries). The requirement of eligibility for naturalization is crucial for the analyses, since it leads to the samples of naturalized and non-naturalized migrants being similar in terms of some individual characteristics, which may be responsible for differences in treatment in the labor markets. The foremost of these is the experience and duration of residency in the destination countries. Since information about the years of residence is missing from most other international and even national datasets that have a sufficient number of immigrants, I consider LIS data at least more appropriate for my analysis compared to other datasets with free or limited access.

Because of the absence of working hours from the LIS database, the individual hourly earnings could not be calculated. For this reason, I assume similar working hours between the groups in each country and calculate the group average hourly earnings using the data on average work hours in the two countries[14]. I use the term "hourly" when discussing the mean decomposition results, but the results are identical to using the total individual employment earnings instead of hourly employment earnings. The largest weakness of the data for the analysis was that the origin countries of immigrants were aggregated to only 7 broad categories for France and were completely missing for Denmark for the chosen year of 2010. Consequently, it was impossible to control for the origins of immigrants and composition of immigrant groups in the two countries.

[12] A more detailed description of the citizenship policies of both countries is in Appendix 2.
[13] The required duration of residency is 5 years in France and 9 years (8 years for refugees) in Denmark. The requirements in Denmark also include renouncement of previous citizenship, passing language and citizenship tests, and being self-supporting.

[14] The average annual working hours was obtained from the economic research website of the Federal Reserve Bank of St. Louis (1,546 in Denmark and 1,478 hours in France).

2.4 Results

I start the analyses by returning to the possible selectivity issues and the conditional independence assumption mentioned in the methodology section. Table 2.1 describes the composition of the data by the employment status of individuals belonging to each group, based on their country, gender, and citizenship status.

The differences in employment rates suggest (but do not indicate) the presence of a possible employment selectivity (selection to employment) issue among migrants. Except for the group of naturalized male migrants in France, all other migrant groups seem to be positively selected into employment based on the low percentages of them being employed (compared to the samples of the native population). Moreover, the groups of eligible migrants in all cases show lower percentages of employed individuals compared to the groups of naturalized migrants, suggesting a higher level of selection to employment for eligible migrants.

Based on the data, around 35% of male migrants and 30% of female migrants who are eligible for naturalization become naturalized in France, whereas in Denmark around half of migrants (belonging to either gender) who are eligible for naturalization become naturalized. This suggests (but again does not indicate) higher possible selection to naturalization in France compared to Denmark.

Both employment selection and naturalization selection issues may create problems for the analyses only if the unobservable characteristics that affect either the work decision or the naturalization decision are correlated with the unobservable characteristics that affect the employment wages and are not correlated with the observable characteristics that I control for. In particular, when estimating the returns to individual characteristics through wage-related regression analysis, such selection biases may cause overestimation of the effects of individual characteristics that have a positive impact on employment or naturalization decisions, and underestimation of the effects of individual characteristics that have a negative impact on such decisions. However, there are several reasons to support the utilization of the conditional independence assumption (CIA) which allows us to neglect such selectivity biases.

First, employment selectivity and naturalization selectivity work in opposite directions (since the data shows that non-naturalized migrants are more selected to employment) and optimistically they may offset each other. Secondly, both the possible employment selectivity and possible naturalization selectivity may create concerns only in the estimations of naturalization premiums when comparing the naturalized migrants to non-naturalized migrants. In estimations of the levels of immigrant integration, I compare both types of immigrants to the randomly-sampled native population. This means that even in the presence of selectivity bias, it may only have a negative impact on the wage gaps and, thus, I will obtain at least the lower bound of differential treatment. In other words, the presence of significant "unobservable" components in native-migrant wage gaps will verify the presence of differential treatment even in the presence of any selection bias.

Table 2.1: Employment Status of Individuals of Working Age by Country, Gender and Type of Citizenship Registration.

			(1)	(2)	(3)	(4)
Country	**Gender**	**Group**	Employed	Un-employed	Not in LF/ Home-maker	N
Denmark	Male	Eligible Migrants	62.65%	6.78%	30.57%	929
		Naturalized Migrants	64.55%	6.36%	29.09%	849
		Natives	84.70%	2.99%	12.32%	29,367
	Female	Eligible Migrants	51.94%	5.68%	42.38%	1,161
		Naturalized Migrants	61.12%	4.82%	34.06%	1,016
		Natives	81.29%	2.56%	16.16%	30,031
Country	**Gender**	**Group**	Employed	Un-employed	Not in LF/ Home-maker	N
France	Male	Eligible Migrants	61.25%	26.84%	11.91%	529
		Naturalized Migrants	83.51%	11.23%	5.26%	285
		Natives	87.69%	8.91%	3.39%	5,475
	Female	Eligible Migrants	27.08%	28.59%	44.33%	864
		Naturalized Migrants	59.95%	16.08%	23.98%	367
		Natives	74.12%	13.02%	12.86%	6,398

Finally, the results of the Oaxaca-Ransom decomposition are almost identical to the results of OLS regressions with a set of control variables in robustness check analyses. This also supports the use of the CIA because unobservable characteristics may impact the outcome variable only through their correlation with the observable characteristics. Further, since the Oaxaca-Ransom model is identical to OLS regression with a set of dummy variables interacted with all explanatory

(observable) variables, one would expect measurable differences in the results in the presence of a selection bias.

The logarithms of the hourly net employment wages (adjusted for purchasing power parity)[15] were employed as the outcome variable in the main regression analyses. The average values of the hourly net employment wages for each group are provided in Table 2.2 (column 2) along with the average values of the main set of explanatory variables (columns 3-7).

Table 2.2: Descriptive Statistics of Average Individual Characteristics of Workers in the Paid-employment Sector by Country, Gender, and Type of Citizenship Status.

			(1)	(2)	(3)	(4)	(5)	(6)	(7)
Country	**Gender**	Person Status	Annual Wages	PPP-adjusted Hourly Wage	Age	Number of Children	Married	Education (3-level rating)	N
Denmark	**Male**	Eligible Migrant	377,969	31.51	40.92	1.34	0.68	2.1	301
		Naturalized Migrant	360,493	30.05	39.34	1.32	0.64	2.26	320
		Native	441,349	36.79	40.63	1.11	0.57	2.23	16,047
	Female	Eligible Migrant	295,674	24.65	40.03	1.58	0.72	2.09	417
		Naturalized Migrant	292,602	24.39	39.31	1.44	0.64	2.29	422
		Native	331,570	27.64	40.73	1.25	0.6	2.35	17,469
France	**Male**	Eligible Migrant	17,988	14.15	40.47	1.8	0.51	1.58	235
		Naturalized Migrant	24,339	19.15	43.02	1.63	0.63	1.96	189
		Native	26,110	20.54	40.84	1.38	0.54	2.1	4,040
	Female	Eligible Migrant	12,806	10.08	39.86	1.81	0.52	1.59	192
		Naturalized Migrant	17,785	13.99	42.48	1.69	0.64	1.96	188
		Native	19,862	15.63	40.75	1.32	0.49	2.22	4,297

Unfortunately, when limiting the data to the samples of individuals employed in the paid employment sector[16] and with non-missing values for all explanatory variables that were used in

[15] This adjustment was introduced to enable me to make comparisons between the results from the two countries. The PPP conversion factors for France and Denmark were obtained from the World Bank official website.

[16] I study only the paid employment sector because of the low number of observations of naturalized individuals employed in the self-employment sector, and the possibility that the wage structure of self-

the regression models, the numbers of observations for each group reduces significantly. Nevertheless, I am convinced that they are still sufficient to ensure the robustness of the results.

I notice some measurable differences in mean hourly wages between the three citizenship-status groups (Table 2.2), which are especially high between the native population and eligible migrants. The native population groups have the highest hourly wages among the three citizenship groups for both countries and genders (column 2), which may indicate citizenship-based differential treatment. However, some measurable differences are also observed in the variables that are generally considered (and proven) to have an impact on the wages, which brings uncertainty as to the source of this inequality. In particular, the groups of native population are prone, on average, to have a higher level of education (column 6) and fewer children (column 4). Both these variables are generally considered to have an impact on hourly wages, which highlights the need for decomposition analysis.

Before proceeding to the decomposition analysis, I first estimated multinomial logit models[17] in order to separately determine occupational attainment probabilities for each citizenship group. I then used these model estimates to predict the occupational allocation probabilities for individuals with mean individual characteristics of each citizenship group, when treated similarly to individuals belonging to another citizenship group. The results of these analyses based on mean individual characteristics are presented in figures 2.1 and 2.2. I divided the occupations into three general groups: managers and professionals, other skilled workers, and laborers and unskilled workers. The latter became the reference group whose coefficients are normalized to zero in the multinomial logit models. The independent variables in these models included age, education, marital status, and the number of children[18].

The differences between the observed and predicted occupational distributions for each citizenship group may indicate occupational segregation in the labor market for the corresponding

employed individuals may be very different from that of individuals employed in paid employment jobs (Hamilton, 2000).

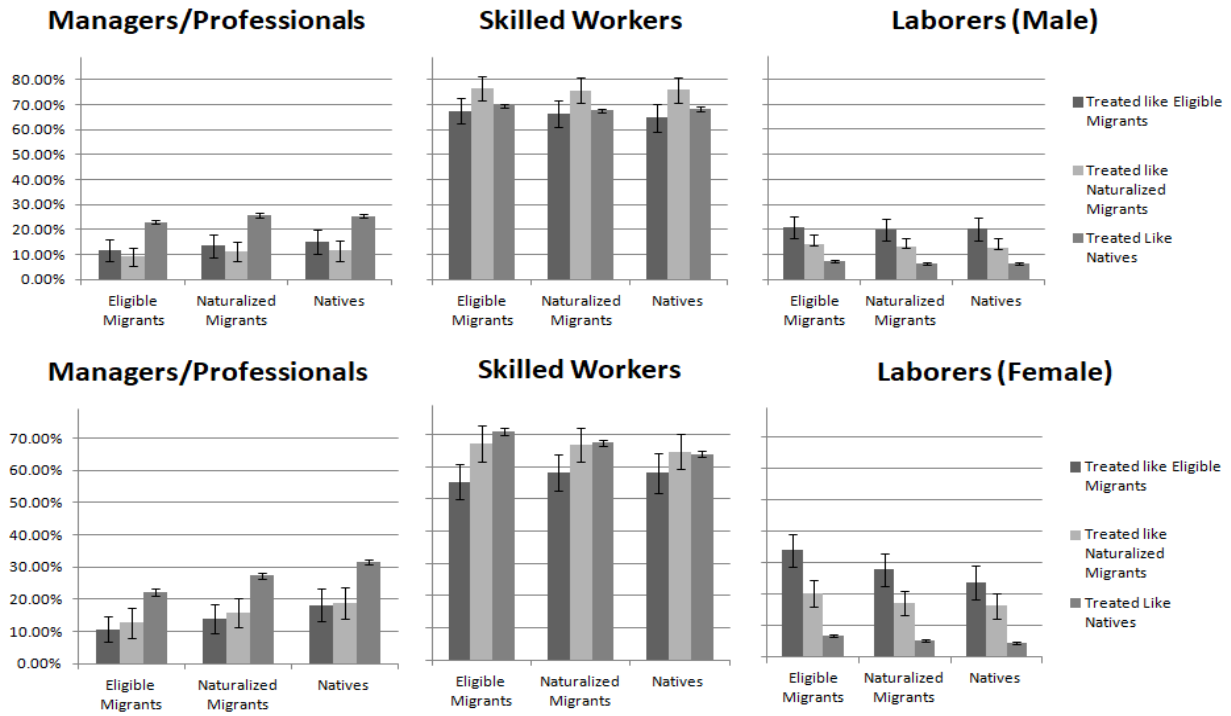
[17] A multinomial logit model was preferred over an ordered probit model because of the absence of obvious hierarchical order among the occupations.

[18] The regression results of the multinomial logit models for each group are presented in tables 7 and 8 in Appendix 1.

country and gender. It is visible from the figures that, based on the observable individual characteristics mentioned earlier, the groups of migrants (both naturalized and eligible for naturalization) are mostly being disadvantageously treated in their occupational allocations. However, there are some specifications that can be highlighted.

In Denmark, a country with stricter naturalization policies, occupational segregation is not only visible in the managerial levels, but it also clearly does not limit itself to the type of citizenship status, since it persists and is mostly concentrated in the columns of native-naturalized migrants' differences in occupational attainment probabilities. Interestingly, the graphs also show possible segregation in the skilled worker positions, but this result is significant only for the female native-eligible migrants column.

Figure 2.1: Multinomial Logit Predictions of Occupational Attainment Models for Denmark for Individuals with Mean Individual Characteristics of Each Citizenship Group by Gender.



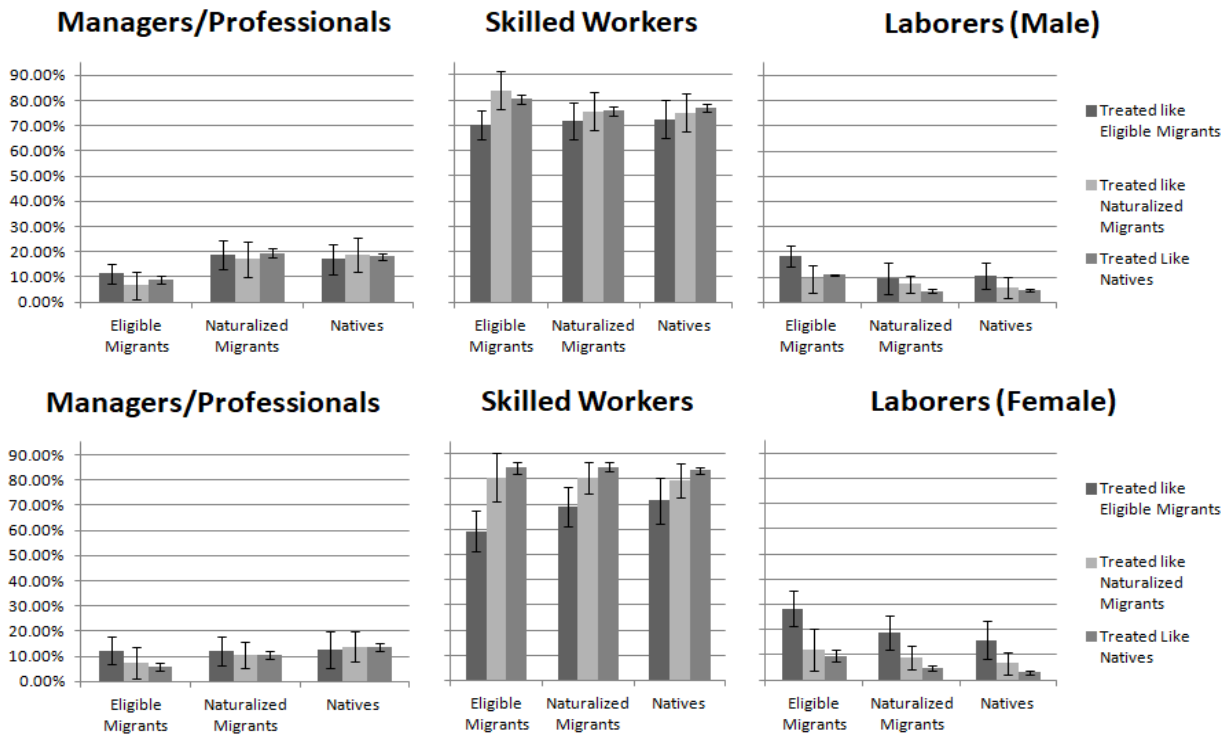
Upper and lower bars represent 95% confidence interval

On the other hand, in France, which has softer naturalization policies, there is almost no occupational segregation at the management level. Moreover, most of the segregation in the skilled positions compared to elementary positions is concentrated between the two types of migrants,

suggesting that the occupational segregation in France at this occupational level is limited mostly to the citizenship status rather than nationality. However, this may also be a result of the possible high naturalization selectivity in France rather than the citizenship status itself.

The first two steps of the decomposition analysis were the separate estimation of the wage structures for each of the 12 groups and the estimation of the "fair" wage structure for both countries and genders. For the sake of concision, I present only the "fair" wage structures in Table 2.3. However, the regression results for each of the 12 groups are presented in tables 2.9 and 2.10 in Appendix 1.

Figure 2.2: Multinomial Logit Predictions of Occupational Attainment Models for France for Individuals with Mean Individual Characteristics of Each Citizenship Group by Gender.



Upper and lower bars represent 95% confidence interval

Table 2.3 shows that the wage structures of individuals are very different depending on the country and gender, which supports the choice of gender-based separation of individuals for the proposed analyses. The statistical significance of chosen independent variables and the corresponding R squared values suggest the robustness of the chosen wage structure identification models.

Table 2.3. OLS Regression Results for the Joint Samples with All 3 Citizenship Status Possibilities by Country and Gender.

VARIABLES	(1)	(2)	(3)	(4)
	Male		Female	
	Denmark	France	Denmark	France
Age	0.088***	0.060***	0.13***	0.043***
	(0.005)	(0.010)	(0.005)	(0.013)
Age Squared	-0.00091***	-0.00058***	-0.0014***	-0.00035*
	(0.000)	(0.000)	(0.000)	(0.000)
Number of Children	0.040***	-0.011	-0.025***	-0.074***
	(0.004)	(0.006)	(0.004)	(0.009)
Married	0.081***	0.11***	0.024**	-0.021
	(0.009)	(0.017)	(0.007)	(0.020)
Firm Size_Middle	0.066***	-	0.11***	-
	(0.011)	-	(0.010)	-
Firm Size_Large	0.17***	-	0.17***	-
	(0.008)	-	(0.008)	-
Supervisor	-	0.17***	-	0.28***
	-	(0.016)	-	(0.022)
Education_Middle	0.13***	0.18***	0.10***	0.22***
	(0.011)	(0.019)	(0.011)	(0.027)
Education_High	0.23***	0.40***	0.17***	0.52***
	(0.013)	(0.024)	(0.013)	(0.032)
Occupation_Manager	0.45***	0.55***	0.47***	0.62***
	(0.016)	(0.033)	(0.016)	(0.050)
Occupation_Skilled	0.20***	0.22***	0.26***	0.26***
	(0.014)	(0.026)	(0.014)	(0.043)
Constant	0.85***	0.97***	-0.16	1.06***
	(0.097)	(0.200)	(0.090)	(0.250)
Observations	17313	4606	19077	4779
R Squared	0.254	0.299	0.219	0.24

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Finally, Table 2.4.1 shows the decomposition results of the average wage differences into the two components that may and may not be explained through the differences in observable characteristics ("explained" and "unexplained" components). Note that all the regression models for wage structures include the occupational statuses of individuals in the same way as presented in Table 2.3. Thus, the differences in average group wages as a result of occupational segregations suggested earlier are being captured in the "explained" components of the wage differences. It is

clear from Table 2.4.1 that the wage differences of native - eligible migrants wage differences are measurably higher in France compared to Denmark. However, it is also visible that very large portions of these differences in France are captured in the explainable component of the columns, indicating that the eligible migrants in France differ considerably from the native population in terms of their observable characteristics.

Table 2.4.1: Hourly Employment Wages Decomposition Results by Country and Gender.

	Denmark Paid-employment (Male)			Denmark Paid-employment (Female)		
VARIABLES	Native - Eligible Migrant	Native - Naturalized Migrant	Naturalized - Eligible Migrant	Native - Eligible Migrant	Native - Naturalized Migrant	Naturalized - Eligible Migrant
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted average earnings of group 1	3.46	3.46	3.23	3.22	3.22	3.06
	(0.00)	(0.00)	(0.04)	(0.00)	(0.00)	(0.03)
Prediction average earnings of group 2	3.31	3.23	3.31	3.03	3.06	3.03
	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Difference	0.15***	0.24***	-0.087	0.19***	0.16***	0.031
	(0.03)	(0.04)	(0.05)	(0.03)	(0.03)	(0.05)
Explained	0.011	0.012	-0.00071	0.097***	0.059***	0.038*
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)
Unexplained	0.14***	0.22***	-0.086	0.092**	0.100***	-0.0076
	(0.03)	(0.04)	(0.05)	(0.03)	(0.03)	(0.04)
Observations	16,348	16,367	621	17,886	17,891	839

	France Paid-employment (Male)			France Paid-employment (Female)		
VARIABLES	Native - Eligible Migrant	Native - Naturalized Migrant	Naturalized - Eligible Migrant	Native - Eligible Migrant	Native - Naturalized Migrant	Naturalized - Eligible Migrant
	(1)	(2)	(3)	(4)	(5)	(6)
Prediction average earnings of group 1	2.87	2.87	2.79	2.55	2.55	2.35
	(0.01)	(0.01)	(0.04)	(0.01)	(0.01)	(0.07)
Prediction average earnings of group 2	2.44	2.79	2.44	1.94	2.35	1.94
	(0.05)	(0.04)	(0.05)	(0.07)	(0.07)	(0.07)
Difference	0.43***	0.074	0.35***	0.62***	0.20**	0.42***
	(0.05)	(0.04)	(0.06)	(0.07)	(0.07)	(0.10)
Explained	0.21***	0.014	0.20***	0.34***	0.11***	0.23***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)
Unexplained	0.21***	0.061	0.15**	0.28***	0.095	0.19*
	(0.04)	(0.04)	(0.06)	(0.06)	(0.06)	(0.09)
Observations	4,275	4,229	424	4,489	4,485	380

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The decomposition results for Denmark (columns 1-3 of Table 2.4.1) indicate that the native-migrant differences in employment wages that are not attributable to individual characteristics ("unexplained" components) are not bound only to the citizenship status of migrants, since such wage differences are not only present but entirely concentrated in the columns of native-naturalized migrants' wage differences for both genders. In contrast to the previous empirical literature (Helgertz, Bevelander and Tegunimataka, 2014), the results suggest that there are no significant naturalization premiums in Denmark for both genders working in the paid employment sector once I compare the migrants to the native population, controlling for the differences in observable characteristics.

Conversely, the decomposition results for France (columns 4-6 of Table 2.4.1) show that the native – migrant differences in hourly wages not attributable to individual characteristics almost entirely depend on citizenship acquisition and mostly disappear after naturalization. Around 70% of the "unexplained" native - eligible migrant wage gaps for both genders are concentrated in the naturalized - eligible migrant columns, with the remaining 30% being statistically insignificant (the "unexplained" components in the native – naturalized migrant columns). These results not only suggest the existence of high naturalization premiums for both genders in France, but also indicate that these naturalization premiums are sufficient for the labor market integration of migrants, making them equable to the native population.

2.5 Robustness Checks

I started the robustness analysis with simple OLS regression models (the results of which are in Table 2.5 and Table 2.6), and proceed with the decomposition of the "unexplained" component of the wage differences to better understand the composition of the wage gaps.

There are two main advantages of using decomposition analysis for this study compared to the OLS models. Firstly, through decomposition analysis it is possible to allow different wage structures for each group and estimate the part of income differences caused by the differences in the returns to characteristics (thus, treatment related). Secondly, I use the wage structures of the general population (described in Table 2.3) as the reference groups when estimating the group differences in wage structures, which allows us to better control for the group characteristics.

Columns with odd numbers in tables 2.5 and 2.6 show the regression results conducted on the combined groups of naturalized and non-naturalized migrants, and are intended to illustrate the naturalization premiums. Columns with even numbers show the regression results conducted on the combined groups of all three citizenship-based sub-samples and are intended to illustrate the differential treatment towards migrants (or migrant integration levels in the destination countries). Columns (1), (2), (5) and (6) in tables 2.5 and 2.6 do not include any other explanatory variables except for the citizenship status of individuals, while columns (3), (4), (7) and (8) include a set of other control variables. Similar to the previous findings, the results suggest no naturalization premiums in Denmark but a significant presence of such premiums in France, which, however, decrease by around half when I include the set of other control variables. I note that this may be a result of the possible naturalization selectivity in France rather than the naturalization itself.

When compared to the native population (the even columns in the tables), the results suggest almost no differential treatment of any kind (in terms of the hourly wages) between immigrants and the natives in France, but this was not the case for Denmark. However, based on the results, the non-naturalized migrants still experience at least some degree of differential treatment in both countries compared to the treatment towards the native population. It is worth mentioning again that because of the possible employment and naturalization selectivity among migrants, these levels of differential treatment may show only the lower bound of the actual differential treatment (or the higher bound of the labor-market integration).

It is also important to notice that the "Difference" rows in Table 2.4 understandably show the same values as the results of the OLS regression models (columns (1), (2), (5) and (6) in tables 2.5 and 2.6), while the rows of the "Unexplained" wage differences in Table 2.4 show values very similar to the corresponding results of the regression models with controls in tables 2.5 and 2.6 (columns (3), (4), (7) and (8)). However, there are some slight differences both in the obtained values and in the levels of statistical significance of the results. In case of any mismatch in the results, the values from the decomposition analysis are considered more reliable due to the two main reasons mentioned earlier (i.e. the possibility of having different wage structures for each group, and "country-equalized" wage structures of the general population).

Table 2.5. OLS Regression Results for Denmark by Type of Citizenship Status and Gender.

Denmark	Male				Female			
VARIABLES	Naturalization Premium	Integration	Naturalization Premium	Integration	Naturalization Premium	Integration	Naturalization Premium	Integration
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Naturalized Migrant	-0.087*	-0.24***	-0.10**	-0.23***	0.031	-0.16***	0.0019	-0.10***
	(0.037)	(0.022)	(0.035)	(0.019)	(0.032)	(0.018)	(0.030)	(0.016)
Eligible Migrant	-	-0.15***	-	-0.14***	-	-0.19***	-	-0.098***
	-	(0.023)	-	(0.020)	-	(0.018)	-	(0.016)
Age	-	-	0.057*	0.086***	-	-	0.094***	0.13***
	-	-	(0.023)	(0.004)	-	-	(0.022)	(0.003)
Age Squared	-	-	-0.00061*	-0.00089***	-	-	-0.0010***	-0.0014***
	-	-	(0.000)	(0.000)	-	-	(0.000)	(0.000)
Number of Children	-	-	0.038*	0.040***	-	-	-0.015	-0.024***
	-	-	(0.015)	(0.003)	-	-	(0.014)	(0.003)
Married	-	-	0.14***	0.086***	-	-	0.069*	0.029***
	-	-	(0.040)	(0.006)	-	-	(0.034)	(0.005)
Firm Size_Middle	-	-	0.15**	0.070***	-	-	0.14**	0.11***
	-	-	(0.054)	(0.008)	-	-	(0.047)	(0.007)
Firm Size_Large	-	-	0.19***	0.17***	-	-	0.18***	0.17***
	-	-	(0.041)	(0.006)	-	-	(0.037)	(0.006)
Education_Middle	-	-	0.084	0.13***	-	-	0.23***	0.11***
	-	-	(0.047)	(0.008)	-	-	(0.043)	(0.008)
Education_High	-	-	0.23***	0.23***	-	-	0.31***	0.18***
	-	-	(0.054)	(0.009)	-	-	(0.050)	(0.009)
Occupation_Manager	-	-	0.46***	0.43***	-	-	0.32***	0.44***
	-	-	(0.060)	(0.011)	-	-	(0.053)	(0.012)
Occupation_Skilled	-	-	0.20***	0.19***	-	-	0.045	0.24***
	-	-	(0.049)	(0.010)	-	-	(0.041)	(0.011)
Constant	3.31***	3.46***	1.41**	0.91***	3.03***	3.22***	0.49	-0.12
	(0.027)	(0.003)	(0.440)	(0.069)	(0.023)	(0.003)	(0.410)	(0.064)
Observations	621	16668	621	16668	839	18308	839	18308
R Squared	0.00436	0.00463	0.193	0.258	0.000547	0.00524	0.159	0.221

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.6. OLS Regression Results for France by Type of Citizenship Status and Gender.

France	Male				Female			
VARIABLES	Naturalization Premium	Integration	Naturalization Premium	Integration	Naturalization Premium	Integration	Naturalization Premium	Integration
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Naturalized Migrant	0.35***	-0.074	0.14*	-0.064	0.42***	-0.20***	0.23*	-0.10*
	(0.064)	(0.044)	(0.060)	(0.038)	(0.096)	(0.056)	(0.096)	(0.050)
Eligible Migrant	-	-0.43***	-	-0.23***	-	-0.62***	-	-0.31***
	-	(0.040)	-	(0.035)	-	(0.055)	-	(0.051)
Age	-	-	0.066	0.060***	-	-	-0.044	0.040**
	-	-	(0.040)	(0.010)	-	-	(0.062)	(0.013)
Age Squared	-	-	-0.00063	-0.00059***	-	-	0.00065	-0.00032
	-	-	(0.000)	(0.000)	-	-	(0.001)	(0.000)
Number of Children	-	-	-0.058**	-0.0093	-	-	-0.094**	-0.069***
	-	-	(0.020)	(0.007)	-	-	(0.036)	(0.009)
Married	-	-	0.27***	0.11***	-	-	-0.047	-0.021
	-	-	(0.062)	(0.017)	-	-	(0.091)	(0.020)
Supervisor	-	-	0.14*	0.18***	-	-	0.31**	0.27***
	-	-	(0.063)	(0.016)	-	-	(0.110)	(0.022)
Education_Middle	-	-	0.14*	0.15***	-	-	0.24*	0.19***
	-	-	(0.069)	(0.020)	-	-	(0.110)	(0.028)
Education_High	-	-	0.25**	0.38***	-	-	0.43**	0.50***
	-	-	(0.089)	(0.025)	-	-	(0.140)	(0.032)
Occupation_Manager	-	-	0.58***	0.52***	-	-	0.57**	0.58***
	-	-	(0.120)	(0.034)	-	-	(0.200)	(0.051)
Occupation_Skilled	-	-	0.21**	0.20***	-	-	0.15	0.22***
	-	-	(0.076)	(0.027)	-	-	(0.120)	(0.044)
Constant	2.44***	2.87***	0.48	0.87***	1.94***	2.55***	2.48*	0.93***
	(0.043)	(0.009)	(0.790)	(0.200)	(0.068)	(0.011)	(1.210)	(0.260)
Observations	424	4464	424	4464	380	4677	380	4677
R Squared	0.0664	0.0255	0.289	0.3	0.0472	0.028	0.213	0.246

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.4.2. Decomposition Results of the Unexplained Component of the Income Differences to the Advantage of the First Group and the Disadvantage of the Second Group.

	Paid-employment (Male)			Paid-employment (Female)		
	Native - Eligible Migrant	Native - Naturalized Migrant	Naturalized - Eligible Migrant	Native - Eligible Migrant	Native - Naturalized Migrant	Naturalized - Eligible Migrant
VARIABLES	(1)	(2)	(4)	(5)	(6)	(7)
Denmark						
Unexplained_1	0.011***	0.011***	-0.21***	0.0090***	0.0090***	-0.091***
	(0.00)	(0.00)	(0.03)	(0.00)	(0.00)	(0.02)
Unexplained_2	0.13***	0.21***	0.13***	0.083***	0.091***	0.083***
	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	16,348	16,367	621	17,886	17,891	839
France						
Unexplained_1	0.017***	0.017***	-0.044	0.017***	0.017***	-0.077
	(0.00)	(0.00)	(0.04)	(0.00)	(0.00)	(0.06)
Unexplained_2	0.20***	0.044	0.20***	0.26***	0.077	0.26***
	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)	(0.06)
Observations	4,275	4,229	424	4,489	4,485	380

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Again, all these results are robust in the presence of the conditional independence assumption (CIA), which was introduced and supported in the beginning of this section. Moreover, when I further decompose the "unexplained" components of the wage differences to the advantages of the first group and the disadvantages of the second group (Table 2.4.2), it becomes more evident that the "unexplained" components obtained in the naturalized - eligible migrant columns for France are less the result of selection to naturalization but rather the naturalization itself, since the groups of naturalized migrants do not have any significant unexplained advantage over the general population.

2.6 Conclusion

In this study, I conduct an employment wage decomposition analysis for naturalized immigrants, non-naturalized immigrants, and the native population at destination for two countries with

different naturalization policies, and compare the returns to the average personal characteristics of individuals belonging to each of the three groups.

The results suggest the existence of high naturalization premiums in France, but do not show any naturalization premiums in Denmark. Moreover, based on the "unexplained" components in the native-migrant employment wage differences, I conclude that, after naturalization, migrants become integrated into the French labor market and are treated similarly to natives. However, this is not the case in Denmark. Importantly, because of the possible employment selectivity issue for migrants as well as other possible selectivity issues for naturalization, which I did not control for in these analyses, the results obtained only reveal the lower bound of the wage differences. Thus, the "unexplained" components in wage differences may be even higher if I generalize the results to the entire sample of migrants. Considering the fact that Denmark has stricter naturalization policies, these findings question the assumption that its policies result in better socioeconomic integration of migrants, and suggest that it may be the socio-cultural attitude to immigrants that results in the formation and strictness level of naturalization policies.

Further studies are still necessary to determine the impact of citizenship policies on the level of socioeconomic integration of migrants. The findings of such studies may help policymakers in the sphere of migration to draft appropriate citizenship polices in order to attain the desired level of immigrant integration in labor markets.

2.A.1 Appendix 1

Table 2.7. Multinomial Logit Regression Results for Occupational Attainment Probabilities for Denmark by Type of Citizenship Status and Gender.

	(1)	(2)	(3)	(4)	(5)	(6)
	Male			**Female**		
VARIABLES	Eligible Migrant	Naturalized Migrant	Native	Eligible Migrant	Naturalized Migrant	Native
Managers						
Age	-0.066	-0.21	0.065	-0.21	0.19	-0.14**
	(0.26)	(0.26)	(0.04)	(0.26)	(0.22)	(0.05)
Age Squared	0.00066	0.00	-0.00065	0.0027	-0.0027	0.0018**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Number of Children	0.027	0.03	0.098**	-0.28*	-0.04	0.13***
	(0.15)	(0.16)	(0.03)	(0.14)	(0.17)	(0.04)
Married	0.19	-0.13	0.29***	-0.19	0.021	0.23**
	(0.40)	(0.47)	(0.07)	(0.37)	(0.37)	(0.08)
Middle Education	1.98**	0.66	1.72***	1.19*	0.82	1.64***
	(0.67)	(0.67)	(0.08)	(0.57)	(0.63)	(0.10)
High Education	3.94***	4.39***	4.89***	4.19***	4.19***	5.28***
	(0.68)	(0.68)	(0.12)	(0.52)	(0.61)	(0.13)
Constant	-1.08	3.72	-2.76***	1.75	-4.86	1.34
	(5.17)	(4.94)	(0.78)	(5.12)	(4.27)	(0.93)
Skilled Workers						
Age	-0.022	-0.16	-0.035	-0.36	-0.039	-0.074
	(0.18)	(0.21)	(0.04)	(0.19)	(0.18)	(0.04)
Age Squared	0.0004	0.00	0.00043	0.0039	0.00011	0.00088
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Number of Children	0.25*	0.06	0.041	-0.17	0.023	0.029
	(0.12)	(0.12)	(0.03)	(0.10)	(0.13)	(0.04)
Married	-0.29	-0.14	0.15*	-0.59*	0.003	0.18**
	(0.30)	(0.36)	(0.06)	(0.28)	(0.30)	(0.07)
Middle Education	0.54	0.35	1.16***	1.20***	1.05***	1.32***
	(0.29)	(0.33)	(0.06)	(0.25)	(0.31)	(0.07)
High Education	0.42	0.98*	2.09***	0.90**	1.05**	1.82***
	(0.37)	(0.46)	(0.11)	(0.31)	(0.39)	(0.12)
Constant	0.86	4.79	1.71*	8.31*	1.85	2.68**
	(3.56)	(3.98)	(0.67)	(3.73)	(3.37)	(0.84)
Observations	453	456	21859	523	549	22370

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.8. Multinomial Logit Regression Results for Occupational Attainment Probabilities for France by Type of Citizenship Status and Gender.

	(1)	(2)	(3)	(4)	(5)	(6)
	Male			**Female**		
VARIABLES	Eligible Migrant	Naturalized Migrant	Native	Eligible Migrant	Naturalized Migrant	Native
Managers						
Age	0.17	0.41	-0.037	-0.014	0.42	-0.09
	(0.30)	(0.41)	(0.09)	(0.30)	(0.50)	(0.12)
Age Squared	-0.0016	-0.01	0.0012	0.00039	-0.0052	0.0015
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Number of Children	0.052	-0.30	0.089	0.09	-0.52*	0.0037
	(0.14)	(0.21)	(0.06)	(0.15)	(0.25)	(0.08)
Married	0.69	-0.56	0.48**	-0.12	-0.38	0.23
	(0.44)	(0.68)	(0.15)	(0.46)	(0.62)	(0.17)
Middle Education	0.6	1.72*	1.69***	-0.46	0.015	1.37***
	(0.55)	(0.79)	(0.18)	(0.84)	(0.77)	(0.23)
High Education	4.44***	4.86***	5.95***	3.38***	2.56**	5.82***
	(1.08)	(1.22)	(0.37)	(0.83)	(0.93)	(0.43)
Constant	-5.58	-8.37	-2.15	-1.31	-7.02	-0.025
	(5.97)	(8.38)	(1.79)	(5.88)	(10.20)	(2.24)
Skilled Workers						
Age	-0.047	0.39	-0.084	0.34	-0.061	-0.16
	(0.21)	(0.32)	(0.08)	(0.22)	(0.42)	(0.10)
Age Squared	0.00089	-0.01	0.0012	-0.004	0.00015	0.0021
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Number of Children	0.088	-0.21	0.067	0.032	-0.29	-0.0048
	(0.10)	(0.15)	(0.05)	(0.11)	(0.18)	(0.07)
Married	0.28	-0.26	0.34**	0.13	0.18	0.17
	(0.30)	(0.58)	(0.13)	(0.33)	(0.51)	(0.15)
Middle Education	0.67	0.01	0.97***	0.84*	0.38	1.02***
	(0.37)	(0.50)	(0.12)	(0.42)	(0.55)	(0.15)
High Education	2.43*	1.50	2.86***	2.03**	0.81	3.15***
	(1.04)	(1.08)	(0.35)	(0.77)	(0.84)	(0.40)
Constant	1.03	-3.92	2.52	-6.55	4.65	4.77*
	(4.03)	(6.54)	(1.50)	(4.22)	(8.49)	(2.03)
Observations	323	237	4789	234	220	4723

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.9. OLS Regression Results for Oaxaca Decomposition for Denmark by Type of Citizenship Status and Gender.

Denmark	Male			Female		
VARIABLES	Eligible Migrants	Naturalized Migrants	Natives	Eligible Migrants	Naturalized Migrants	Natives
Age	-0.0083	0.075*	0.087***	0.13***	0.080**	0.13***
	(0.032)	(0.032)	(0.004)	(0.034)	(0.028)	(0.003)
Age Squared	0.00018	-0.00086*	-0.00090***	-0.0014***	-0.00085*	-0.0014***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Number of Children	0.068**	0.011	0.040***	0.014	-0.052*	-0.025***
	(0.021)	(0.021)	(0.003)	(0.020)	(0.021)	(0.003)
Married	-0.052	0.29***	0.084***	0.0049	0.14**	0.027***
	-0.052	-0.06	-0.0063	-0.051	-0.047	-0.0052
Firm Size_Middle	-0.0085	0.30***	0.068***	0.18*	0.13*	0.11***
	(0.071)	(0.080)	(0.008)	(0.071)	(0.063)	(0.007)
Firm Size_Large	0.04	0.31***	0.17***	0.16**	0.22***	0.16***
	(0.053)	(0.063)	(0.006)	(0.056)	(0.050)	(0.006)
Education_Middle	0.13*	0.069	0.13***	0.27***	0.17**	0.098***
	(0.060)	(0.072)	(0.008)	(0.058)	(0.064)	(0.008)
Education_High	0.33***	0.15	0.23***	0.30***	0.27***	0.17***
	(0.071)	(0.080)	(0.009)	(0.070)	(0.073)	(0.009)
Occupation_Manager	0.41***	0.52***	0.43***	0.40***	0.25**	0.46***
	(0.074)	(0.094)	(0.012)	(0.075)	(0.077)	(0.012)
Occupation_Skilled	0.21***	0.18*	0.19***	0.032	0.045	0.26***
	(0.060)	(0.077)	(0.010)	(0.057)	(0.061)	(0.011)
Constant	2.89***	0.87	0.89***	-0.19	0.82	-0.17*
	(0.640)	(0.620)	(0.070)	(0.670)	(0.530)	(0.064)
Observations	301	320	16047	417	422	17469
R Squared	0.183	0.24	0.258	0.179	0.151	0.223

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.10. OLS Regression Results for Oaxaca Decomposition for France by Type of Citizenship Status and Gender.

France	Male			Female		
VARIABLES	Eligible Migrants	Naturalized Migrants	Natives	Eligible Migrants	Naturalized Migrants	Natives
Age	0.075	0.0099	0.060***	-0.0069	-0.071	0.046***
	(0.055)	(0.058)	(0.011)	(0.092)	(0.089)	(0.013)
Age Squared	-0.00064	-0.000074	-0.00059***	0.00017	0.00098	-0.00039*
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Number of Children	-0.084**	-0.008	-0.0022	-0.09	-0.12*	-0.066***
	(0.027)	(0.030)	(0.007)	(0.050)	(0.053)	(0.010)
Married	0.34***	0.15	0.091***	-0.14	0.075	-0.02
	(0.086)	(0.090)	(0.017)	(0.130)	(0.130)	(0.021)
Supervisor	0.14	0.089	0.18***	0.53**	0.1	0.27***
	(0.093)	(0.084)	(0.016)	(0.170)	(0.140)	(0.022)
Education_Middle	0.20*	0.082	0.15***	0.23	0.25	0.19***
	(0.098)	(0.094)	(0.021)	(0.160)	(0.150)	(0.029)
Education_High	0.14	0.36**	0.39***	0.2	0.64***	0.50***
	(0.140)	(0.110)	(0.026)	(0.210)	(0.180)	(0.033)
Occupation_Manager	0.67***	0.50**	0.51***	0.76*	0.23	0.60***
	(0.190)	(0.160)	(0.036)	(0.300)	(0.300)	(0.055)
Occupation_Skilled	0.20*	0.19	0.20***	0.25	-0.09	0.25***
	(0.098)	(0.130)	(0.030)	(0.150)	(0.220)	(0.049)
Constant	0.15	2.02	0.88***	1.75	3.43	0.78**
	(1.090)	(1.160)	(0.210)	(1.760)	(1.770)	(0.260)
Observations	235	189	4040	192	188	4297
R Squared	0.284	0.235	0.291	0.186	0.207	0.233

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2.A.2 Appendix 2

2.A.2.1 Citizenship Legislation in Denmark

Danish citizenship laws were created in 1950 and have been gradually tightened ever since, establishing one of the most stringent citizenship policies in Europe. As of 2010, the requirements for naturalization included residency in Denmark for 9 years (8 years for refugees), renouncement of previous citizenship, and passing a language and a citizenship test. Moreover, the individuals must also swear an oath of allegiance and loyalty to Denmark and be self-supporting, i.e. not receiving any social benefits during the year prior to naturalization. Since, in my dataset, I do not know whether or not an individual is a refugee, I classified the migrants as eligible for naturalization if they have lived in Denmark for at least 9 years and did not receive any social benefits on the 9^{th} year of residency. The foreign-born share of the population in Denmark is around 7-8%, the majority of whom are immigrants from Asia (40-45%), with the remainder mainly originating from Eastern Europe, Africa, and Latin America.

2.A.2.2 Citizenship Legislation in France

France has one of the softest and most accessible citizenship policies among EU member countries. It has tolerant policies for dual citizenship and the required conditions for citizenship include residency of 5 years (effective since 1945). The requirements also include other conditions, such as no criminal conviction of a certain type and being a person of good character, that are generally standard in most other European countries. During the period of 1990-2010 the country implemented four citizenship legislation changes (1993, 1998, 2003, and 2006), but none of these changes was fundamental. In my analysis, an immigrant in France was categorized as eligible for naturalization if he lived in France for at least 5 years. Immigrants constitute around 13% of the French population, with 4.5% being naturalized. The majority of immigrants in France originate from Eastern Europe (40-45%) and Africa (30-35%), and the remainder is mainly from Asia and Latin America.

Chapter 3: Social Integration of Immigrants and the Attitude of the Native Population in European Countries

3.1 Introduction

The successful social integration of immigrants and their children in the destination countries is crucial for many reasons. Most importantly, successful integration of immigrants and better social acceptance of immigrants by the host societies result in a higher contribution to the host country. Conversely, poor integration may result in the exclusion of immigrants and their children from the social and economic norms of the destination countries. In both cases, the attitude of the native population may serve as both the cause and the result of the degree of immigrant integration, and, in turn, may impact the immigration and naturalization policies in the host countries.

Integration of immigrants into host country economies and societies has become a popular topic for research and discussion, particularly in recent decades. However, much integration-related research focuses on the socioeconomic integration of immigrants expressed through the health, education, and labor market outcomes of immigrants in the destination countries. Another stream of research addresses only the attitude of the native population towards immigrants and immigration in the main destination countries. Little work to date has studied the level of immigrant integration outside the context of the labor market, or in combination with the attitude of the native population towards immigrants and immigration.

The notion of "naturalization" is defined in this study as the acquisition of citizenship of the destination country by immigrants. This definition is widely used in the literature and results in very clear limits and plain differentiation between naturalized and non-naturalized immigrants (Bevelander and Veenman, 2008; Engdahl, 2011). Since naturalization is the procedure that results in the elimination of all (or almost all) legal barriers and differences between immigrants and natives, it is reasonable to consider it when studying the level of the social integration of immigrants.

In this study, I focus on the social integration of immigrants and the perceived discrimination of immigrants in host societies. The notion of social integration of immigrants is

discussed in this chapter through the evaluation of a feeling of closeness to the host country and a measure of the frequency by which the official language of the host country is being spoken at home. The latter can be considered an indicator of acculturation, which is defined as the convergence of the behavior of immigrants to that of the native population. The perceived discrimination of immigrants is evaluated through their answers to a survey question as to whether or not they consider themselves a member of a group that is being discriminated against in the country.

As the main contribution to the existing literature, I attempt to reveal the relationships between both the social integration and perceived discrimination of immigrants, and the attitude of the native population towards immigrants and immigration (similar to related literature, the ATII abbreviation is used hereafter). As a highly complex measure, ATII can clearly both impact immigrant integration, and be impacted by immigrant integration. For this reason, different indicators of ATII are employed in this study to cover its different aspects. These indicators include the willingness of the native population to accept new immigrants into their homeland, the natives' opinion about the role of the government in immigration-related issues, and their opinion about the contribution of immigrants to their country. Moreover, the impact of ATII on immigrant integration may also be indirect, through effecting institutionally-imposed restrictions and integration-related policies. In order to estimate the direct effect of ATII, institutionally-imposed restrictions in the form of the stringency of naturalization policies are also controlled for in this study.

The objective of this study is to provide a more comprehensive picture of the integration of immigrants in European countries by first analyzing whether naturalization policies reflect ATII, and then attempting to link ATII and institutionally-imposed restrictions in the form of naturalization policies with the social integration of immigrants and the perceived discrimination of immigrants in these countries. The empirical study is conducted on 20 European countries using similar norms and definitions that could allow cross-country comparison.

3.2 Literature Review

The academic literature related to this study is relatively extensive, covering three core aspects: natives' attitude towards immigrants and immigration, naturalization policies, the integration of immigrants in host societies, as well as their interactions. In the past decade, a large body of

research has been developed on these classic issues. However, most of these studies focus on only one of the three aspects, while a collective presentation could give a more complete picture of the causes and determinants of immigrant integration.

The natives' attitude towards immigrants is usually covered in social studies and sociological research. Ceobanu and Escandell (2010) present a good critical review of this literature, which also covers European countries, and attempt to understand the causes and implications of ATII. The authors highlight the importance of and need for cross-national research that would link ATII to the institutional environment and the immigrant composition in the receiving societies. They conclude that, because of the vast complexity of the issue, the area lacks theoretically and empirically extended research.

Nevertheless, some research studies have attempted to reveal the relationship between ATII and macroeconomic or other country-specific characteristics of host countries, mostly related to outcomes in labor markets. For example, using longitudinal data, Gorinas and Pytlikova (2017) investigate whether anti-immigrant attitudes affect migration flows in OECD countries, and show that the natives' tendency to discriminate in the labor market has a robust negative effect on migration flows, particularly for migrants from more developed countries. Other studies have attempted to reveal the relationships of ATII with the unemployment rates in receiving countries (Meuleman, Davidov and Billiet, 2009), the presence and strength of radical right-wing parties in receiving countries (Bohman and Hjerm, 2016; Gorinas and Pytlikova, 2017), and native-immigrant differences in skills and labor market competitiveness (Mayda, 2006).

Other recent theoretical and empirical research (mainly developed by Facchini, and Mayda) attempts to relate ATII with the immigration policies in the host countries. For example, some studies analyse not only the determinants of ATII from theoretical and empirical perspectives, but also how these attitudes are reflected in migration policies (Facchini and Mayda, 2008; Facchini and Mayda, 2009; Facchini and Mayda, 2012; Hatton, 2017). The authors provide some evidence that while restrictive migration policies are in line with the predictions of the median-voter framework, there are still large gaps between ATII and the actual policies, which may be explained by the salience, pressure groups, lobbyists, and policymakers' concerns about social welfare. However, to my knowledge, no study has attempted to reveal to what extent ATII is reflected in a country's naturalization policies.

Most integration-related studies also focus mainly on the socioeconomic integration of immigrants in labor markets, ignoring other sides of integration. Such studies, conducted for separate countries, generally find employment and wage gaps between natives and immigrants (including naturalized immigrants), even after controlling for observable individual heterogeneity (Steinhardt, 2012; Fougere and Safi, 2008; Bevelander and Veenman, 2008). More comprehensive studies include the analysis of the level of labor market integration by immigrants in 29 European countries by Dustmann and Frattini (2011). Using data from the European Labor Force Survey (EULFS), the authors show that in most European countries immigrants are highly disadvantaged in the labor markets, even when their performance is compared to that of the native population with the same measurable skills. Such findings suggest that unsuccessful integration of immigrants in labor markets may be a cause of poor social integration, which has not been properly studied in the literature.

Moreover, there is very little literature that examines the relationship of the host countries' naturalization policies with the socioeconomic integration of immigrants. One such study is Gathmann and Keller (2014), who examine two major immigration reforms in Germany to test whether more liberal access to citizenship can improve the economic integration of immigrants. The results suggest that liberalization of citizenship access provides some benefits in the labor market but does not completely eliminate the gaps between immigrants and the native population.

There are other cross-national studies related to the impact of citizenship policies on the socioeconomic integration of immigrants. For example, Vink, Prokic-Breuer and Dronkers (2013) conduct a cross-sectional study of 16 European countries and show that more accessible citizenship policies have a significant impact on naturalization rates for immigrants from less developed countries. However, in their study the authors do not differentiate between immigrant integration and the rates of naturalization.

Finally, using data from the European Social Survey (ESS) from 2001 to 2009, Aleksynska and Algan (2010) study both the cultural and economic integration of first and second generation immigrants in European destinations, and show that integration-related economic and social outcomes may take place at different speeds and do not have systematic correlations. The authors also relate the destination country's integration policies to the assimilation of immigrants and show that policies that favor labor market integration also favor immigrant assimilation. However, the main focus of the study was to check correlations between

different outcomes among native-born and first and second generation immigrants rather than reveal the connections between ATII, the naturalization policies, and the social integration of immigrants in the countries of residency.

This study builds on and contributes to the present literature by creating a more comprehensive picture of immigrant integration by simultaneously covering ATII, its relation to a country's naturalization policies, as well as the social integration and perceived discrimination of immigrants.

3.3 Data

The main analyses are conducted using the 2010-2014 data waves (2010, 2012 and 2014) of the European Social Survey (ESS). Due to its suitability and the wide variety of variables, this survey is often used by researchers in evaluating the public opinion of the native population towards immigrants and immigration (Aleksynska and Algan, 2010; Hatton, 2014; Hatton, 2017). The start wave of 2010 was chosen because of the 2009 implementation of the "Treaty of Lisbon", an important agreement that provides a consolidated legal foundation for the EU to promote and support the immigrant integration practices of member states. The end wave of 2014 was chosen since it was the last ESS data wave at the time of the research and in order to exclude the possible impact of the recent migration crisis in Europe. Among the advantages of the ESS is the recurrence of the same questions with the same definitions across survey waves and countries, which makes it possible to conduct cross-national comparative analyses. As of 2014, ESS includes complete or partial data on 23[19] of the 28 EU member states. However, Croatia (HR) was removed from the analysis since it joined the EU in 2013 and only had a data wave from 2010. Two more countries (Bulgaria and Hungary) were also excluded from the analysis for other reasons explained later in this section.

Using various indicators of the ESS, I was able to measure most aspects of the social and cultural integration of immigrants, as well as the attitude of the native population towards immigrants and immigration. These indicators include the feeling of closeness to the host country and language acquisition (which I refer to as "social integration"), the level of perceived

[19]The 23 EU member state ISO country codes in ESS are: AT, BE, BG, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IT, LT, NL, PL, PT, SE, SI, SK.

discrimination, as well as the native population's willingness to allow more immigrants into the country and their opinion about the contribution of immigrants (which are referred to as ATII).

In addition, the survey contains such key indicator variables as the birthplace of individuals and their parents, and the duration of their residency in the host country. This allows us to identify an immigrant as a person who was born in a country other than the country of current residence (country where the survey was taken) and to further differentiate the individuals into four groups:

1) Non-naturalized immigrants - citizens of countries other than the country of residence and not born in the country of residence;
2) Naturalized immigrants - citizens of the country of residence not born in the country of residence;
3) Second generation immigrants- citizens of the country of residence that were born in the country of residence and have at least one parent not born in the country of residence[20];
4) Native population - citizens of the country of residence that were born in the country of residence with both parents born in the country of residence.

This differentiation of individuals was conducted to be able to estimate any changes in the attitude of the native population and the social integration of immigrants that are related to the citizenship status or the generation of immigrants. When using this differentiation, I also controlled for possible changes of country borders that might raise a concern if the respondents were not certain as to what country code to report as a county of birth. Since 2012, the ESS database includes DDR, USSR, Czechoslovakia, Yugoslavia, East Timor, Serbia and Montenegro in the list of possible birthplaces. Fortunately, there were few observations with such reports, and after their elimination, I concluded that this issue should not create a significant problem for the analysis.

On the other hand, three of the major constraints of the data are: 1) the surveys are conducted only in the country language and may not reach those who do not speak the language

[20] A more widely accepted definition of "second generation immigrant" is to be born to two non-native parents. However, because of the low number of such observations for many countries in the data, this non-standard definition was adopted in the main analyses. Robustness checks with the use of the more widely accepted definition for the countries with an appropriate number of observations showed similar results. These Analyses are available upon request.

(who are potentially less integrated); 2) the small migrant sizes by different origins in the data sample may not be representative of the entire immigrant population; 3) the differences in immigrant composition between countries challenge the robustness of cross country comparisons.

To overcome the first constraint, I conducted additional analyses were conducted with restricted data including only the long-term immigrants who already reside in the destination countries for at least 5 years[21] (and thus are highly likely to speak the language). The results of these analyses with the restricted data were in line with the results of the main analyses.

To overcome the second constraint, I compare the immigrant samples in the ESS data with the 2015 data on the total immigrant population from the EUROSTAT Migration and Migrant Population Statistics Report[22] and exclude those countries from the ESS database that show a significant mismatch in the relative immigrant shares[23]. Based on the comparison of the relative shares of the immigrant population statistics from the ESS database to the EUROSTAT 2017 Migration and Migrant Population Statistics Report (Table 3.1), two more countries (BG and HU) with the highest levels of mismatch were dropped from further analyses, leaving only 20 EU member states. In addition, six more countries (AT, CZ, GR, IT, PL, and SK) require careful consideration and, in some cases, are also dropped from the analyses either because of the mismatch or because of being present only in one data wave[24], which results in a lower number of observations.

Finally, to overcome the third constraint, I use origin-to-host country pair controls in the regression analyses. Moreover, I am extremely careful in making conclusions regarding the causality of the relationships because of the possible reverse causality issues and the possible impact of country specific characteristics on the results.

---

[21] With such restriction, the group of naturalized immigrants also becomes more comparable with the group of non-naturalized long-term immigrants since the minimal required duration of residency for the purposes of naturalization in EU member states is at least 5 years. The results of these analyses are available upon request.

[22]The 2017 EUROSTAT Report was selected since it uses data up to 2015 and was considered the most comparable source of the statistics for this study.

[23]The shares of foreign born population between EUROSTAT and ESS results are clearly not perfectly identical, but are mainly very close to each other. Any mismatch of more than two times of at least two indicators was considered significant for the exclusion from further analysis.

[24] The data from AT is present only in the 2014 wave; GR data is present only in the 2010 wave; and IT data is present only in the 2012 wave. BG and SK do not have the 2012 year waves. All other country data is present in all three waves.

Table 3.1: Share of Foreign Born Population (Immigrants) in Destination Countries by the Country of Residence and the Place of Birth.

Country	ISO	Foreign born population (EUROSTAT)			Foreign born population (ESS)			
		Total	EU born	Non EU born	Total	EU born	Non EU born	N
Austria	AT	18.2%	8.2%	10.0%	11.8%	5.6%	6.1%	211
Belgium	BE	16.3%	7.7%	8.6%	12.7%	5.9%	6.8%	676
Bulgaria	BG	1.9%	0.7%	1.2%	0.7%	0.4%	0.4%	34
Czech Republic	CZ	4.1%	1.6%	2.5%	2.1%	1.8%	0.3%	137
Germany	DE	13.3%	5.3%	8.0%	9.8%	4.2%	5.7%	884
Denmark	DK	11.2%	3.8%	7.4%	7.0%	2.6%	4.4%	332
Estonia	EE	14.7%	1.5%	13.2%	17.1%	0.7%	16.4%	1063
Spain	ES	12.7%	4.2%	8.5%	10.2%	3.0%	7.2%	579
Finland	FI	6.0%	2.2%	3.8%	4.2%	1.8%	2.4%	259
France	FR	11.8%	3.3%	8.5%	10.5%	2.9%	7.6%	584
United Kingdom	GB	13.3%	5.0%	8.3%	12.2%	3.4%	8.8%	850
Greece	GR	11.3%	3.2%	8.1%	9.8%	2.4%	7.4%	267
Hungary	HU	5.1%	3.3%	1.8%	1.8%	1.2%	0.6%	94
Ireland	IE	16.9%	11.6%	5.3%	14.6%	10.4%	4.1%	1101
Italy	IT	9.7%	3.0%	6.7%	7.3%	3.0%	4.3%	70
Lithuania	LT	4.5%	0.7%	3.8%	3.6%	0.4%	3.3%	219
Netherlands	NL	12.1%	3.3%	8.8%	8.8%	2.4%	6.4%	490
Poland	PL	1.6%	0.6%	1.0%	1.2%	0.6%	0.6%	64
Portugal	PT	8.4%	2.2%	6.2%	6.7%	1.1%	5.6%	372
Sweden	SE	17.0%	5.4%	11.6%	12.6%	4.8%	7.8%	643
Slovenia	SI	11.7%	3.2%	8.5%	8.0%	3.0%	5.1%	309
Slovakia	SK	3.3%	2.8%	0.5%	2.2%	1.8%	0.4%	81

Note: EUROSTAT presents the national 2015 European Union statistics on international migration, while the ESS data-based calculations show the shares of the foreign born population residing in the country in 2014.

Table 3.2 presents the 2010-2014 average individual characteristics of each of the four identified groups in each of the 20 EU member states. Because of the low number of observations of non-naturalized immigrants in PL and SK, these two countries were also excluded in corresponding parts of the analysis (related to non-naturalized immigrants).

It can be observed that, in line with logical expectations, the group of naturalized immigrants is often the oldest among the three immigrant groups (except for EE where non-naturalized immigrants are the oldest, and FI and GR[25] where second generation immigrants are

[25] In Greece, the group of second generation immigrants was actually the oldest among the four groups. The high average age of this group is most likely a result of immigration of co-ethnic Greeks from Albania and other Balkan nations following the collapse of the Soviet Union in 1989.

the oldest among the groups of immigrants) but is still on average younger than the group of the native population (except for AT, CZ, IT, LT, PL, SI, and SK).

Table 3.2: Average Individual Characteristics by Group and Country of Residence.

Country	Non-naturalized Immigrants								Naturalized Immigrants							
	Age		Years of Education		Female	Married	Presence of Children	N	Age		Years of Education		Female	Married	Presence of Children	N
	(Mean)	(SD)	(Mean)	(SD)	(%)	(%)	(%)		(Mean)	(SD)	(Mean)	(SD)	(%)	(%)	(%)	
AT	41.0	13.4	12.9	4.1	54%	47%	46%	114	52.8	16.3	12.6	3.7	62%	53%	39%	97
BE	41.4	15.2	12.8	4.4	52%	52%	50%	341	46.6	16.4	12.1	4.3	56%	62%	55%	335
CZ	45.2	15.8	13.6	3.0	35%	50%	50%	26	56.6	16.0	12.6	3.0	60%	50%	35%	111
DE	41.5	13.5	13.0	4.7	58%	66%	52%	338	47.2	18.1	13.3	3.6	50%	63%	41%	545
DK	44.1	16.4	13.7	5.7	55%	58%	44%	159	46.0	16.4	14.0	5.7	57%	56%	46%	173
EE	62.4	14.2	12.0	3.4	62%	57%	34%	658	61.8	15.0	12.5	3.7	74%	52%	24%	405
ES	38.5	13.6	12.6	5.1	49%	49%	42%	382	44.2	14.0	14.4	6.0	54%	54%	54%	196
FI	35.7	12.8	15.1	4.3	50%	35%	46%	125	39.9	15.5	15.1	4.1	52%	40%	43%	134
FR	44.7	15.8	12.3	4.9	56%	55%	46%	234	54.3	16.8	12.6	4.9	54%	52%	40%	350
GB	37.6	14.2	15.2	4.5	57%	46%	37%	316	50.0	17.7	13.5	4.3	52%	59%	47%	533
GR	37.9	11.3	10.9	3.4	58%	61%	58%	178	43.1	17.3	12.5	4.0	60%	56%	40%	88
IE	37.3	12.8	15.0	3.7	48%	46%	41%	697	42.2	14.9	15.1	3.4	57%	49%	46%	403
IT	39.4	12.2	12.8	3.9	50%	53%	59%	34	49.2	16.7	11.9	4.6	60%	54%	51%	35
LT	51.1	15.1	13.5	2.5	57%	48%	22%	23	60.6	13.4	12.4	3.6	58%	46%	32%	196
NL	41.1	12.5	13.6	4.9	48%	41%	42%	132	49.0	15.9	13.7	4.8	58%	43%	38%	358
PL	46.8	18.7	16.4	5.0	40%	80%	60%	5	69.0	16.1	10.3	4.2	59%	48%	40%	58
PT	37.7	12.3	11.2	4.5	62%	36%	40%	149	46.5	15.8	11.0	4.8	60%	43%	50%	223
SE	43.5	16.9	13.6	4.4	48%	49%	45%	157	47.7	17.6	13.3	3.8	55%	47%	39%	485
SI	40.9	14.9	11.3	3.3	38%	74%	60%	47	55.6	13.8	10.9	4.3	54%	71%	60%	262
SK	46.1	17.4	13.2	3.5	36%	55%	64%	11	55.3	16.0	13.4	3.1	67%	63%	44%	70
Total	43.5	16.4	13.2	4.5	54%	52%	43%	4,126	50.1	17.3	13.1	4.4	57%	54%	43%	5,057

Country	Second Generation Immigrants								Native Population							
	Age		Years of Education		Female	Married	Presence of Children	N	Age		Years of Education		Female	Married	Presence of Children	N
	(Mean)	(SD)	(Mean)	(SD)	(%)	(%)	(%)		(Mean)	(SD)	(Mean)	(SD)	(%)	(%)	(%)	
AT	46.3	18.7	13.0	3.5	49%	41%	35%	153	50.0	18.2	12.3	3.1	52%	45%	30%	1,418
BE	40.3	18.4	13.1	4.0	49%	39%	36%	500	48.4	19.3	13.0	3.6	51%	49%	37%	4,085
CZ	49.6	16.7	12.8	2.6	51%	46%	32%	422	46.8	17.3	12.7	2.4	51%	47%	35%	5,935
DE	44.6	17.4	13.9	3.4	53%	47%	27%	728	49.9	18.6	13.8	3.4	49%	54%	30%	7,300
DK	43.8	20.7	13.4	5.0	51%	41%	28%	247	49.1	18.7	13.1	5.1	48%	52%	32%	4,111
EE	45.0	16.1	13.4	3.1	59%	47%	41%	880	48.0	20.1	12.9	3.5	58%	38%	37%	3,907
ES	38.8	18.7	13.7	4.6	55%	32%	34%	102	48.3	18.6	12.5	5.8	50%	53%	44%	5,004
FI	44.3	23.1	13.2	4.6	51%	22%	19%	110	50.7	19.0	13.1	4.4	51%	34%	28%	5,753
FR	46.9	18.7	12.8	3.6	53%	39%	33%	694	51.0	18.7	12.5	4.0	54%	43%	32%	4,299
GB	45.2	17.4	14.2	3.9	58%	39%	35%	563	52.9	18.9	13.1	3.5	56%	47%	30%	5,531
GR	60.9	20.2	9.2	4.4	51%	51%	25%	142	47.8	18.7	11.4	4.3	56%	54%	38%	2,294
IE	40.1	18.6	14.5	3.3	58%	35%	33%	301	49.4	18.5	13.5	3.5	54%	47%	36%	6,149
IT	36.2	18.7	14.1	4.2	44%	28%	32%	25	47.7	18.4	12.6	5.3	52%	49%	39%	855
LT	50.5	17.3	12.8	3.2	60%	46%	36%	395	49.5	19.0	12.4	3.5	62%	45%	33%	5,399
NL	43.8	19.0	13.9	4.1	59%	36%	30%	405	51.8	17.9	13.5	4.1	54%	49%	32%	4,666
PL	54.3	13.4	12.4	3.4	53%	72%	55%	200	45.3	18.9	12.4	3.5	53%	56%	47%	4,972
PT	39.2	20.5	9.9	4.7	58%	25%	27%	146	54.2	19.0	7.7	5.0	59%	54%	33%	5,018
SE	38.6	16.4	13.3	3.0	51%	33%	35%	436	50.2	19.6	12.8	3.5	50%	45%	30%	4,024
SI	42.0	18.5	12.5	3.4	54%	39%	38%	335	48.5	18.9	12.0	3.5	54%	49%	45%	3,191
SK	53.5	16.7	12.8	2.8	64%	52%	45%	156	49.7	17.0	12.9	3.0	60%	55%	44%	3,452
Total	45.1	18.4	13.2	3.7	55%	42%	34%	6,940	49.6	18.8	12.6	4.1	54%	48%	35%	87,363

Note: The statistics are calculated using only one year wave data for AT, GR and IT (2014, 2010 and 2012 respectively), resulting in a lower number of observations. BG and SK do not have 2012 year waves. The rest of the countries have all three data waves and the average statistics are calculated.

In terms of years of education, all four groups stand more or less close to each other, but the groups of naturalized and second generation immigrants often show higher average results than the non-naturalized immigrants or even the natives. The non-naturalized immigrants are often, on average, more likely to be married and to have children than any other group, with the group of naturalized immigrants being the close second in both categories. However, there is no clear group dependent tendency in the share of females within a group.

Finally, to scale the countries based on the level of strictness of their naturalization policies, I employ the Migrant Integration Policy Index (MIPEX), which also allows cross-country comparisons. MIPEX is a unique tool which assembles indices measuring the favorability of citizenship and integration policies for immigrants in European countries. These indices cover 167 indicators of various dimensions in order to describe migrants' inclusion in the society of the host country. The indicators are from 8 main policy dimensions: "Labor Market Mobility", "Education", "Political Participation", "Family Reunion", "Health", "Permanent Residence", "Anti-discrimination", and "Access to Nationality". Since I am interested only in the naturalization polices, the indicators of "Access to Nationality" are the main focus of this research, and the basis on which the countries are classified.

According to the MIPEX documentation, this policy dimension covers four sub-categories that cover 35 indicators in total. Each indicator can take one of the five possible values between 0 and 100 (0, 25, 50, 75 or 100), and the grade of the sub-category is the mathematical average of the indicators included in it. Thus, the average value of this index ranges from 0 to 100, with 0 being the most unfavorable for immigrants and 100 being the most favorable. Table 3.3 reports the sequence of the MIPEX score of "Access to Nationality" for each of the 20 countries for the period 2010-2014 and the five year average values. The table does not reveal many within-country variations in the scores, which suggests the strength of the index and its use as an indicator of the stringency of a country's naturalization policies in this analysis.

The first sub-category of the MIPEX indicator of "Access to Nationality" is "Eligibility", which generally converges on requirements regarding the duration of residence in the host country. The second sub-category is the "Conditions for Acquisition of Status", which covers the requirements regarding language, citizenship and integration tests, income and job security as well as the financial fees for acquiring the citizenship. The third sub-category is the "Security of Status", which covers the entitlement to naturalization, the grounds for rejection, and

possibilities to appeal. Finally, the fourth sub-category is "Dual Nationality" and its legal acceptance by the host country[26].

Table 3.3: MIPEX Scores of "Access to Nationality" by Country and Year.

		MIPEX Score of "Access to Nationality"					
Country	ISO	2010	2011	2012	2013	2014	**Average**
Austria	AT	27	26	26	26	26	26.2
Belgium	BE	62	62	62	69	69	64.8
Czech Republic	CZ	40	40	40	40	49	41.8
Germany	DE	66	66	66	66	72	67.2
Denmark	DK	35	35	35	42	58	41
Estonia	EE	18	18	18	18	18	18
Spain	ES	48	48	48	48	48	48
Finland	FI	61	63	63	63	63	62.6
France	FR	61	61	60	61	61	60.8
United Kingdom	GB	62	62	62	60	60	61.2
Greece	GR	57	57	57	34	34	47.8
Ireland	IE	57	59	59	59	59	58.6
Italy	IT	50	50	50	50	50	50
Lithuania	LT	30	35	35	35	35	34
Netherlands	NL	68	68	68	66	66	67.2
Poland	PL	26	26	56	56	56	44
Portugal	PT	86	86	86	86	86	86
Sweden	SE	73	73	73	73	73	73
Slovenia	SI	41	41	41	41	41	41
Slovakia	SK	25	25	25	25	25	25

Note: The table is constructed using the Migrant Integration Policy Index from http://www.mipex.eu.

3.4 Methodology

As mentioned in the previous section, the measures of various social and cultural indicators of individuals were constructed using answers to ESS survey questions with different scaling systems. The country and group averages of these measures and other descriptive statistics were then computed to estimate the relationships between various indicators. The results of these analyses are presented in section 3.5.

From a purist standpoint, it is technically not correct to take averages of any ordinal scale because ordinal values may carry exactly the same amount of information being scaled using any

[26] The scores of the sub-categories of the MIPEX indicator of "Access to Nationality" for 2010-2014 periods can be found in Table 3.6 in the Appendix. The complete list of included indicators and the covered questions of each sub-category is in Table 3.7 in the Appendix.

other proportional scaling system. In fact, the sensitivity of empirical studies to monotonic transformations of ordinal scales has been highlighted by many researchers (Bond and Lang, 2013; Schröder and Yitzhaki, 2017). However, from a pragmatic point of view, this methodology not only works but is also an accepted practice in almost all social science fields and is accompanied by standards for how well it works. Discussions about controversies regarding the use of traditional descriptive statistics for ordinal-level variables have been going on for decades[27].

The simplified approach to dealing with ordinal variables as if they were numeric requires the assumption that the numerical distance between each set of subsequent categories is equal and justified. If that assumption is very close to reality, then analyses based on these numbers will render veritable results. In this case, any proportional change in the scaling system will clearly change the mean values, but will have no impact on either the correlation coefficients or the significance of the regression models and regression coefficients. The appropriateness of the scaling systems of the ESS dataset may be justified by the ESS documentation. However, readers can decide themselves whether or not they consider the numeric distances close to reality.

Some technically correct alternatives to this simplified approach include the use of the median and mode instead of averages (however, these can sometimes provide poor summaries of the data), the use of appropriate binomial values instead of ordinal scaling systems, and the use of Logit, Probit or other ordinal regression models instead of OLS. All of these methods have been used in suitable parts of the analyses in order to confirm the robustness of the results[28]. In general, my conclusions were identical using different scaling systems and parametric statistics. The current version of the analysis is presented in this chapter in order to be more interpretable and understandable for the reader.

The same complications with applicability of general statistical methods are basically valid for any ordinal-scale indicator, including the MIPEX index and its categories that were also used in this study. However, the practical use of such ordinal indexes is so widely used among academic researchers that there are many published academic articles in related fields that use this index, without even mentioning the issue (Aleksynska and Agan, 2010; Huddleston, Niessen, Ni Chaoimh and White, 2011; Stadlmair, 2017b). Moreover, the very little within-country

[27] Early history about such conflicting views may be found in Gardner, 1975.
[28] The results of the robustness check analyses with other parametric statistics, and analyses with the use of aggregated scaling systems, are available upon request.

variation in the index during the observed 5-year period (Table 3) also supports the use of this index as an aggregated indicator of the stringency of naturalization policies.

In addition to the descriptive results, regression analyses were conducted to better evaluate the relationship between ATII or naturalization policies and the social integration and perceived discrimination of immigrants. Because of some restraints of the data and the methodological restrictions related to ordinal-scaled variables mentioned earlier, the choice of the outcome variables for the regression analyses were restricted to binomial variables (the perceived discrimination and language acquisition of immigrants), and appropriate robustness analyses were conducted to strengthen the obtained results. Moreover, because of possible reverse causality issues (described in more detail in the result section) the regression analyses are aimed at the estimation of the relationships between the variables rather than the causality. For this reason, OLS regression analyses were conducted to evaluate the statistical significance of the regression coefficients.

Since there is no developed theoretical literature that would suggest an empirical model for estimation of the impact of various country-specific or individual characteristics on social integration or perceived discrimination of immigrants, an empirical model was developed based on related literature from the labor economics. In particular, the set of explanatory variables in the regression models of integration was built based on some empirical literature that estimates the impact of naturalization on the labor market performance of immigrants (Steinhardt, 2012; Helgertz, Bevelander and Tegunimataka, 2014; Gathmann and Keller, 2014; Sargsyan, 2017). As a result, in addition to the main explanatory variables of interest (ATII and institutionally imposed restrictions on immigrant naturalization), the empirical model employed in this study includes age, education, duration of residency in the host country, gender, marital status, presence of children, and the citizenship status of immigrants (naturalized vs. non-naturalized).

In line with logical expectations, most of these variables, including the level of education, duration of residency, and naturalization were shown to have a positive impact on the labor market integration of immigrants (Gathmann and Keller, 2014; Sargsyan, 2017). There are no obvious reasons to expect a different outcome when estimating the respective impacts on the social integration or perceived discrimination of immigrants. Moreover, the findings of related empirical literature on the social and cultural integration of immigrants (Mayda, 2006; Aleksynska and Algan, 2010; Keller, Gathmann and Monscheuer, 2015), also suggest using similar controls with similar expectations.

As a result, the following model was estimated to identify the impact of ATII and the institutionally imposed restriction (expressed through the MIPEX indicators) on the social integration of immigrants:

$$Y_i = \beta_0 + \beta_1 R_i + \beta_2 E_i + \beta_3 G_i + \beta_4 M_i + \beta_5 C_i + \beta_6 N_i + \beta_7^n ATII_n + \beta_8^m MIPEX_m + \beta_9^k Other_{ik} + e_i, \quad (3.1)$$

where Y is a variable of interest responsible for integration (perceived discrimination of immigrants, or the language acquisition as an indicator of the social integration) of immigrant i; R is the years of residence in the host country; E the years of education; G, M, C and N are binomial dummy variables standing respectively for the gender, marital status, presence of children in the household, and naturalization of immigrants; ATII is a vector of indicators of the country-average attitude of the native population towards immigrants and immigration (n)[29]; MIPEX is a vector of indicators of the institutionally imposed restriction on immigrants expressed through the MIPEX scores (m); Other is a vector of other characteristics (k) including other individual characteristics, origin-to-host country pair controls and data wave controls; and e is the error term. The regressions were conducted on the joint samples of naturalized and non-naturalized immigrants.

Finally, to cover the impact of naturalization polices in more detail and further restrict the possible bias of the ordinal scale of the MIPEX scores, all the four sub-categories of the indicator of "Access to Nationality" were used as separate explanatory variables: "Eligibility", "Conditions for Acquisition of Status", "Security of Status", and "Dual Nationality"[30]. As mentioned in the Data section, each sub-category can take values from 0 to 100, with 0 being the most unfavorable for immigrants and 100 being the most favorable.

[29] Because of the high correlation between the ATII indicators (see Table 3.8 in the Appendix), these indicators were used separately in the regression models in order to exclude multicollinearity.
[30] The indicators used for the estimation of each of the four MIPEX categories can be found in Table 3.7 in the Appendix.

3.5 Results

The results in this section present the country average levels[31] of the indicators of interest and follow the objective to describe the indicators of ATII, social integration, perceived discrimination and naturalization policies that were used in this study[32] for each of the 20 countries for which the study is conducted. The sub-sections also provide the correlation and regression coefficients between different pairs of these indicators in order to shed light on the respective relationships. However, because of possible reverse causality issues, discussed in detail in the sub-sections, the conclusions based on these findings should be done very carefully. The sub-sections are structured as follows: sub-section 3.5.1 focuses on the naturalization policies and their relationships with the ATII indicators. Sub-section 3.5.2 presents the analysis of the perceived discrimination of immigrants, while sub-section 3.5.3 presents the analysis of the social integration of immigrants.

3.5.1 Naturalization Policies and ATII

In this sub-section, I provide some insight into ATII expressed through the viewpoints of the native population in the host countries regarding new immigrants and their contribution, as well as the role of the government in immigration-related issues. This matter is of great importance since it is the native population's votes and general public opinion that, in theory, form and impact a country's immigration-related policies, including naturalization policies (Facchini and Mayda, 2008; Facchini and Mayda, 2012)[33]. Thus, it is reasonable to expect the causality of this relationship to be from ATII to the formation of the policies. However, I cannot exclude the possibility that the causality of the relationship may also be in the other direction, since the

[31] Similar figures were obtained when using median values instead of averages. The results of these analyses are available upon request.

[32] As mentioned in the Data section, because of the low number of observations, the group average results of non-naturalized immigrants were eliminated for PL and SK. In addition, all the estimates based on the questions that were added to the ESS questionnaire only in year 2014 (the feeling of closeness to the country of residence, the average willingness to allow more immigrants from poorer countries in Europe; native populations' opinions about the contribution of immigrants to crime problems, the treatment of the government towards new immigrants, and government generosity in judging refugee applications) are also absent for countries that did not have this data wave (i.e. GR, IT and SK).

[33] In their studies, the authors discuss the median-voter perspective and interest-group dynamics as two possible mechanisms through which ATII may impact the immigration-related policies. Both of the mechanisms assume the causality of this relationship to go from ATII to the stringency of immigration-related policies, with the second channel having more potential to explain the actual stringency of the policies.

naturalization policies may impact the selectivity of the immigrants which, in turn, may form the natives' opinions. Nevertheless, parallel to the related literature, I hypothesize that the countries where the native populations express negative attitudes towards immigrants and immigration will have less favorable naturalization policies towards immigrants.

The attitudes of the native population are expressed in this study through averaging the natives' answers to some survey questions regarding their willingness to accept new immigrants of specific origins into their country, their desired level of stringency in immigration-related policies, and their opinion about the contribution of immigrants to the country of residence.

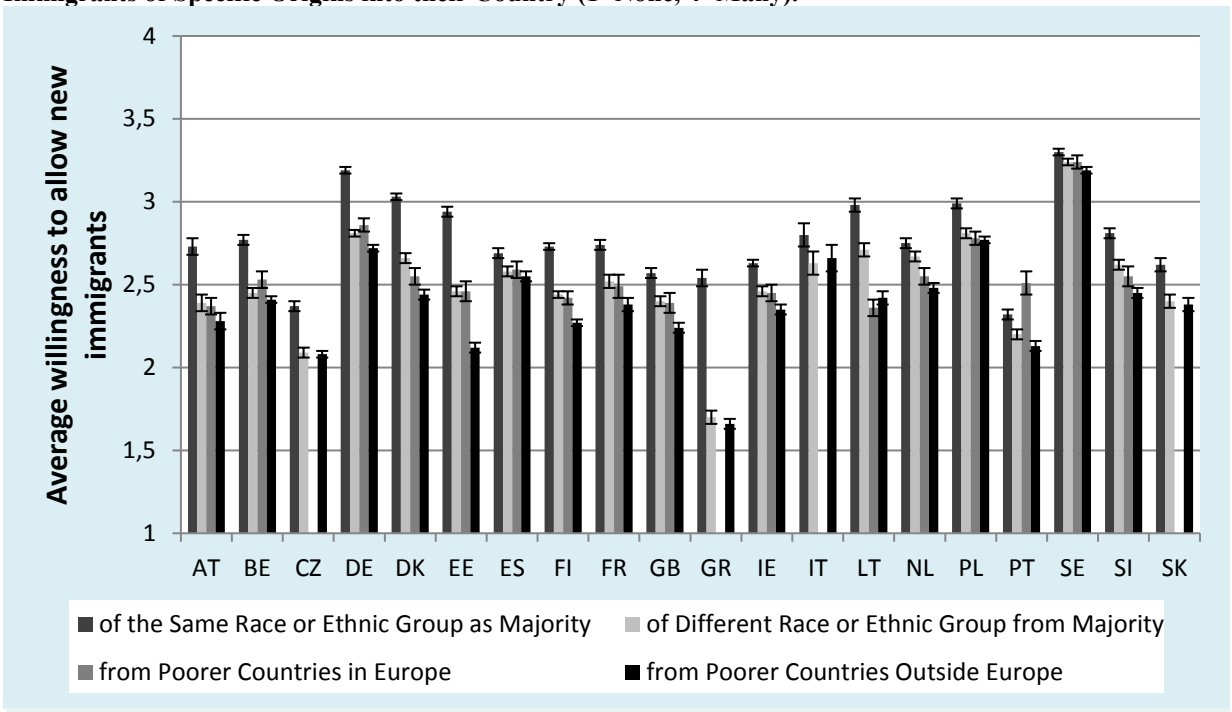
Figure 3.1.1 shows the country-average willingness of the native population to allow new immigrants of specific origins into their country. The scaling of the willingness indicators varies from the lowest score of 1 for low willingness to allow new immigrants of specific origin, to the highest score of 4 for willingness to allow many immigrants. In general, it is observable that within-country variation of each indicator is rather small. This supports the robustness of the indicators and the use of country-averages as country representative values.

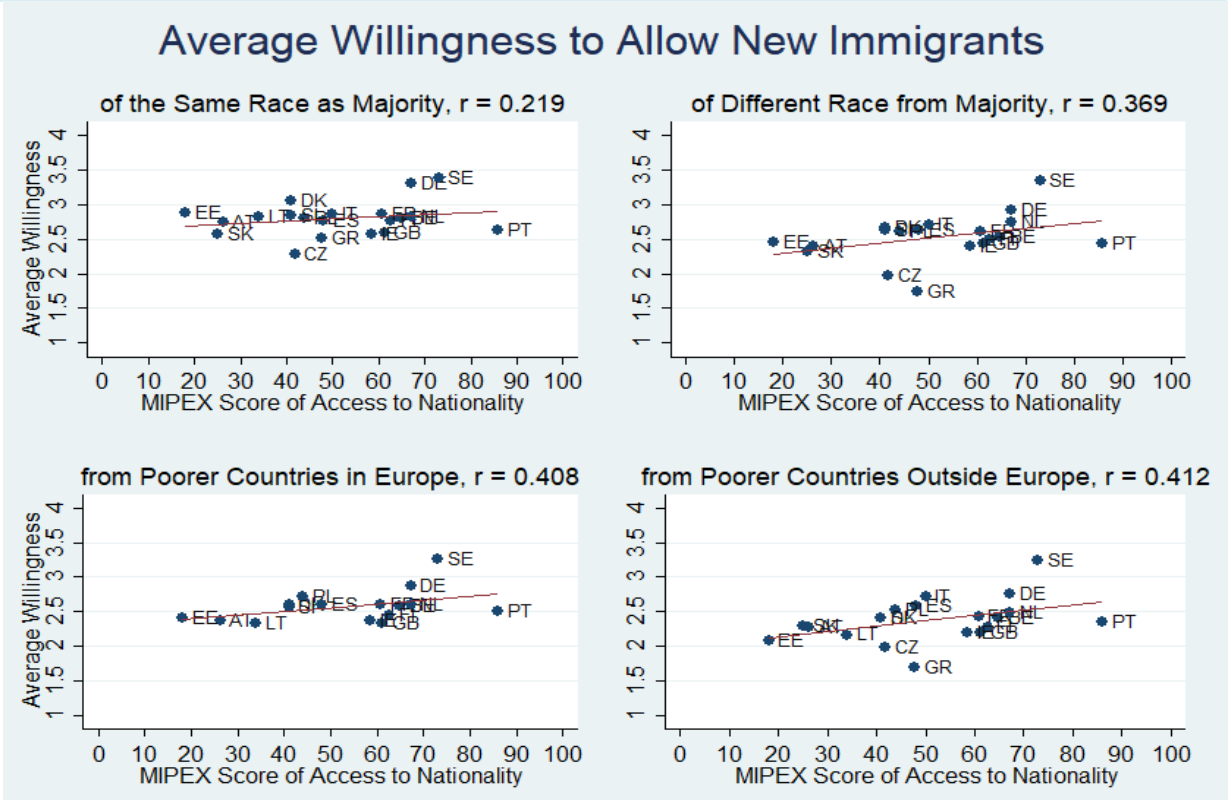
The upper part of the figure shows that the country-average willingness levels to allow more immigrants from poorer countries outside Europe is lower than that for immigrants of any other classification in all 20 countries, varying from around 1.7 for Greece to around 3.2 for Sweden. The latter shows the most favorable relation towards new immigrants of all origins expressed by the native population, while the former shows the most unfavorable relation.

With the exception of Portugal, the willingness to allow new immigrants of the same race or ethnic group as the majority is always higher than that for migrants from other origins or ethnic belonging. Portugal may actually be considered as an outlier in most cases, having below average levels of willingness of the native population to accept new immigrants with the most favorable naturalization policies for immigrants.

The lower part of Figure 3.1.1 presents the distribution of country-average statistics based on the MIPEX index score of "Access to Nationality". The figure shows that the native population's average level of willingness to allow new immigrants into their country is positively related with the MIPEX score and thus is fairly reflected in the naturalization policies of the country.

Figure 3.1.1: The Average Willingness of the Native Population in Destination Countries to Allow New Immigrants of Specific Origins into their Country (1=None, 4=Many).





The upper and lower bars stand for a 95% confidence interval, r is the correlation coefficient.

Note: The survey questions used for the creation of these indicators for each of the 4 specified options (of the same race as the majority; of different race from the majority; from poorer countries in Europe; and from poorer countries outside Europe) were: "Would you like to allow many/few immigrants of 'the specified group' to the country?" The scaling of these survey questions could take values {1, 2, 3, or 4}, ranging from 1 for "few" to 4 for "many". The values of the MIPEX scores range from 0 (the most unfavorable for immigrants) to 100 (the most favorable for immigrants).

Based on the correlation coefficients, the average level of willingness to allow new immigrants from poorer origins into their country is a better predictor of the country's naturalization policies, than the average willingness to allow new immigrants of specific ethnic belonging. This finding suggests that the stringency of a country's naturalization policies is more related with the desire of the native population to restrict access of immigrants from poorer countries.

I proceed with the native populations' opinion about the role and attention of the government towards immigrants. Figure 3.1.2 presents the natives' average opinions about the relative treatment towards new immigrants by the government (hereafter treatment towards new immigrants) and the generosity of the government in judging applications for refugee status (hereafter generosity in judging refugee applications) for each country. The results do not show much between-country variation of these indicators, with the average values being concentrated between the scores of 2.5 and 3.5. In about half of the observed countries, the native population on average believes that new immigrants are being treated better than themselves, although there is not much statistical or economic significance in the level of these differences.

There are only five countries (DE, DK, FI, SE and SI) where the native population on average believes that immigrants are being treated worse than themselves. Two of these countries (DK and SI) have below average MIPEX scores, while the other three (DE, FI and SE) have above average MIPEX scores, indicating no clear relationship between the MIPEX score and this indicator. This result is also observable from the lower left graph of the figure.

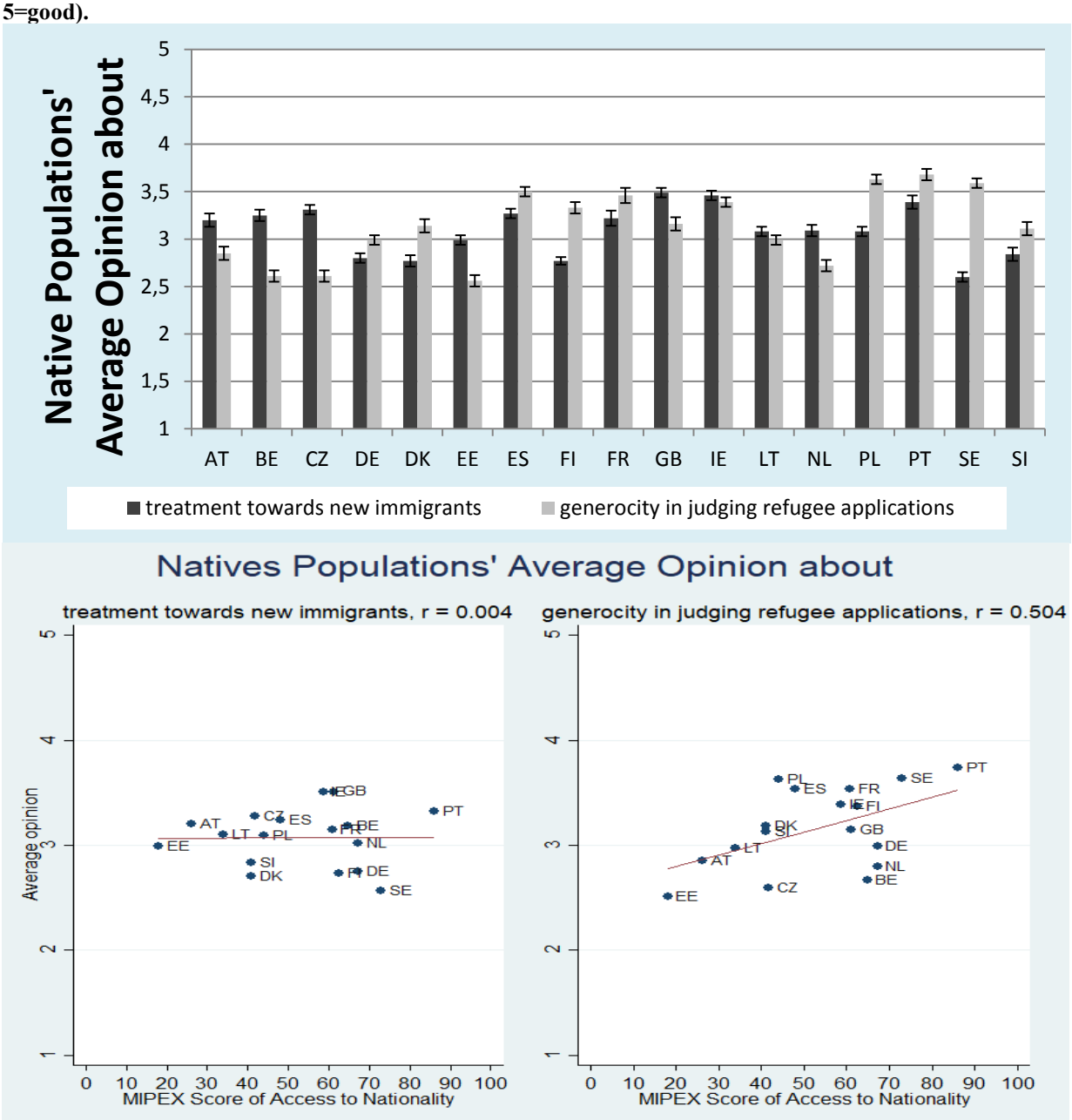
In terms of the native populations' opinion regarding the generosity of governments in judging refugee applications, the figure does not show much between-country variation of the average values, with a high concentration of the scores in the region between 2.5 and 3.5. However, an evident positive correlation with the MIPEX score is visible from the lower right graph of Figure 3.1.2. This finding suggests that the natives' average desired level of generosity in judging refugee applications is also fairly represented in the stringency of naturalization policies.

Finally, Figure 3.1.3 describes the native populations' opinion about the contribution of immigrants to their country, economy, culture, and the crime problems[34]. The scaling in this

[34]Native populations' opinion about the contribution of immigrants to crime problems was estimated through the question: "Immigrants make the country's crime problems worse (0) or better (10)". GR, IT and SK estimates are missing from the figure because only the 2014-year wave contains this question.

figure ranges from 0 for "making the indicator worse" to 10 for "making it better" with 5 standing for "no impact".

Figure 3.1.2: The Native Populations' Average Opinion about the Treatment of the Government Towards New Immigrants and the Generosity of the Government in Judging Applications for Refugee Status (1=bad, 5=good).



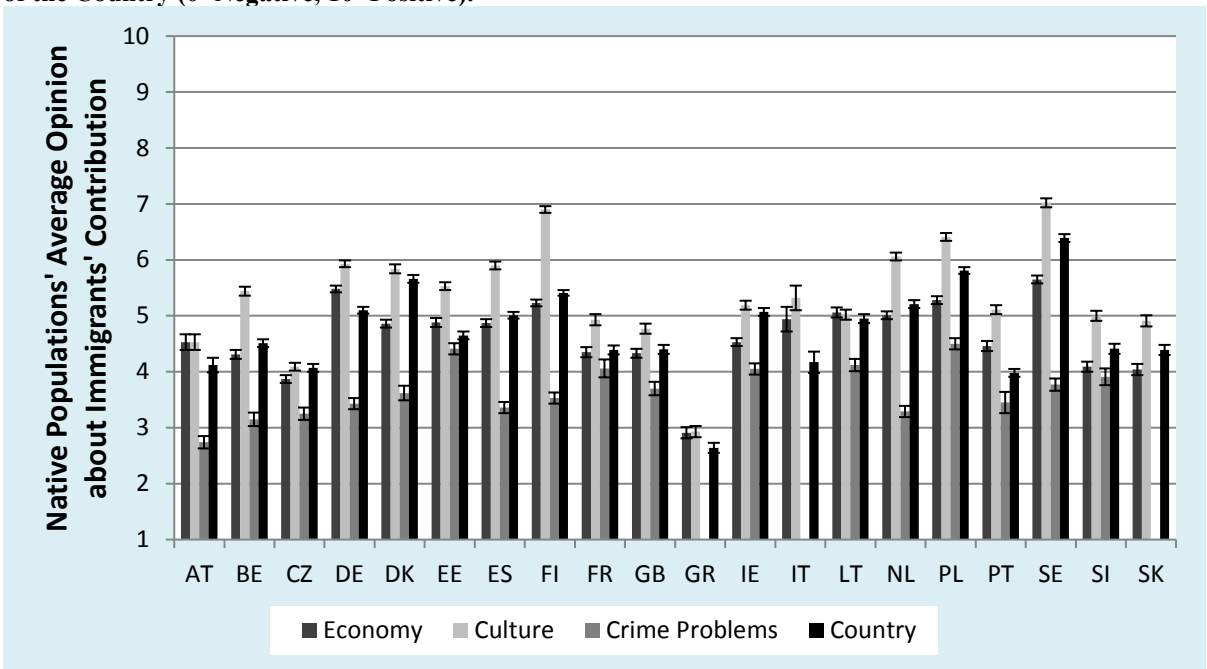
The upper and lower bars stand for a 95% confidence interval, r is the correlation coefficient.

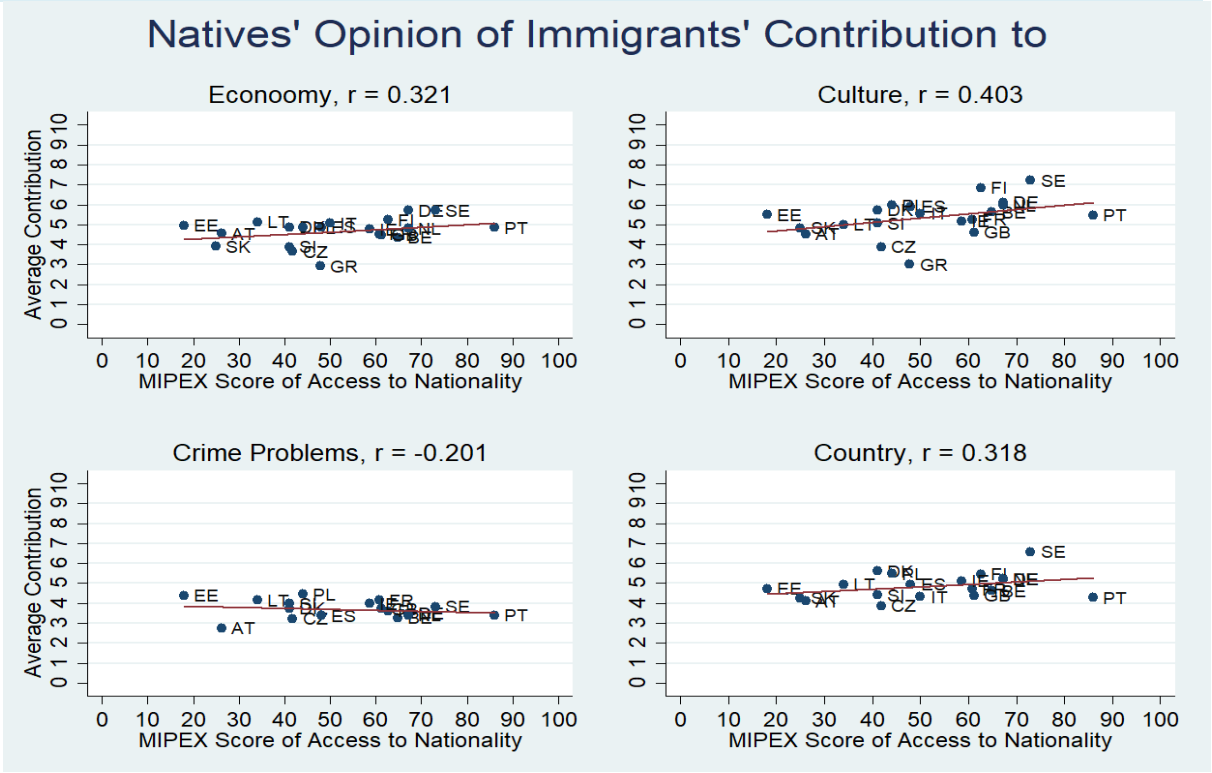
Note: The survey questions used for the creation of these indicators were statement questions: "Compared to yourself, government treats new immigrants better" and "Government should be more generous in judging applications for refugee status". The scaling of these survey questions could take integer values from 1 to 5, ranging from 1 for "strongly disagree" to 5 for "strongly agree". The values of the MIPEX scores range from 0 (the most unfavorable for immigrants) to 100 (the most favorable for immigrants).

Here again, Sweden presents the most positive opinions of the native population regarding the contribution of immigrants, with Finland standing not far behind, while Greece presents the most negative opinions. Figure 3.1.3 shows a high concentration of the country-average statistics

below the average score of 5.0, with the estimates of cultural contribution being the largest in all countries.

Figure 3.1.3: The Native Populations' Average Opinion about Immigrants' Contribution to Different Aspects of the Country (0=Negative, 10=Positive).





The upper and lower bars stand for a 95% confidence interval, r is the correlation coefficient.

Note: The survey questions used for the creation of these indicators for each of the 4 specified options of contributions (to the economy; culture; crime problems; and the country in general) were: "Do you consider immigration to be bad or good for 'the specified aspect' of the country?" The scaling of these survey questions could take integer values from 0 to 10, ranging from 0 for "making the indicator worse" to 10 for "making it better". The values of the MIPEX scores range from 0 (the most unfavorable for immigrants) to 100 (the most favorable for immigrants).

The relation of the MIPEX index score to the opinion of the native population regarding immigrants' contribution to different aspects of the country (lower part of Figure 3.1.3) also reveals some interesting results. However, it should be mentioned that the high concentration of average statistics around the score of 5.0 that stands for "no impact" makes the results less conclusive.

The figure shows a positive correlation between the MIPEX score and the natives' opinion of immigrants' positive contribution to the economy, culture, and the country in general, with the cultural aspect revealing the strongest positive correlation. Interestingly, the correlation with the contribution to the crime problems appears to be negative, indicating that in countries with more favorable naturalization policies for immigrants, the natives believe that crime problems worsen with immigration. However, no conclusions may be drawn about causality, because of the differences in the initial states of the crime indicators between the countries.

The results in this sub-section generally show that countries where the natives express more welcoming and liberal opinions regarding immigrants and immigration have more favorable naturalization policies for immigrants. This finding supports the hypothesis of a positive correlation between ATII and the favorability of naturalization policies for immigrants.

3.5.2 Perceived Discrimination of Immigrants

This sub-section offers insights into the perceived discrimination of immigrants in the host countries and its relation to the ATII and the naturalization policies in those countries. The indicator of perceived discrimination was constructed through individuals' answers to the question of whether they consider themselves a member of a group that is discriminated against in their country of residence. With this in mind, a negative relationship of ATII (or the MIPEX score) with the indicator of perceived discrimination of immigrants would mean lower perceived discrimination of immigrants in countries with more positive ATII (or more favorable naturalization policies for immigrants).

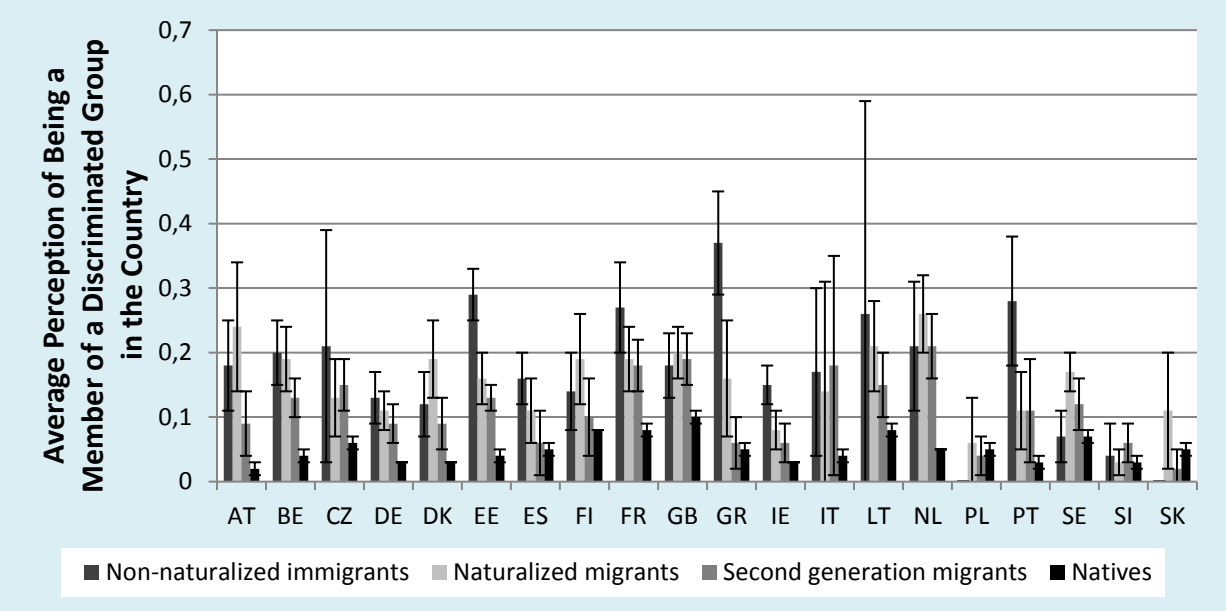
In studies of these relationships, most of the related literature (Facchini and Mayda, 2008; Facchini and Mayda, 2012) simply assumes that the causality of the relationships go from ATII and naturalization policies to the perceived discrimination of immigrants. However, even though it is difficult to imagine a potential issue with the reverse causality in these relationships, this possibility can still be present. Poor initial integration and high perceived discrimination of

immigrant may result in more negative ATII and tighter naturalization policies in order to assure better integration of future immigrants.

Since the possibility of reverse causality cannot be completely excluded, the aim of this analysis is to estimate the sign of the relationship rather than the causality. To that end, I hypothesize that immigrants will express lower perceived discrimination in countries with more positive ATII or more favorable naturalization policies for immigrants. Moreover, because of the positive correlation between the ATII indicators and the MIPEX score of "Access to Nationality" shown in the previous sub-section, the impact of the stringency of naturalization policies on the perceived discrimination of immigrants could rather be considered one of the mechanisms of the indirect impact of ATII.

The average perceived discrimination by each of the four groups (including the natives) in the countries of their residence are described in Figure 3.2.1. One noticeable and expected trend is that the perceived discrimination of second generation migrants is almost exclusively lower than that of first generation naturalized migrants. However, this difference is never statistically or economically significant.

Figure 3.2.1: The Group-Average Perception of Being a Member of a Discriminated Group in the Country of Residence.



The upper and lower bars stand for a 95% confidence interval.

Note: The survey question used for the creation of this indicator was a yes (1) or no (0) question: "Do you consider yourself a member of a group discriminated against in this country?" The group averages of this indicator were separately computed for the groups of non-naturalized immigrants, naturalized immigrants, second generation immigrants, and the native population.

Interestingly, the results show that in some countries (AT, DK, FI, NL and SE), naturalized immigrants report higher average perceived discrimination than non-naturalized immigrants.

However, this difference is economically measurable and statistically significant only for Sweden. On the other hand, measurable differences in the average levels of perceived discrimination in favor of the group of naturalized migrants can be found in Estonia and Greece. Another interesting finding is that the perceived discrimination of the natives is, on average, also non-zero in all countries, and in some cases even reaches measurable levels.

The figure shows very large within-country variations in perceived discrimination of all three groups of immigrants, which highlights the need for individual-level analysis. There are two hypotheses with opposite impacts that I want to test through the evaluation of the perceived discrimination of immigrants. On the one hand, a negative attitude of the native population towards immigrants and stricter naturalization policies could result in tighter selection and better integration of immigrants. In this case, I would expect to see positive relationships between the ATII (or MIPEX) indicators and the perceived discrimination of immigrants.

On the other hand, positive attitudes of the native population towards immigrants and softer naturalization policies could cause a more open social life and better social integration of immigrants, resulting in negative relationships between the ATII (or MIPEX) indicators and the perceived discrimination. This also means that the absence of a visible one-way relationship may be a result of both of these forces working together.

To reveal the relationships between perceived discrimination and the ATII or the MIPEX indicators of the stringency of naturalization policies, regression analyses were conducted on the joint sample of first-generation naturalized and non-naturalized immigrants[35]. Table 3.4 presents the results of these regression analyses of OLS regression models based on equation (3.1). In the base models in column 1, only the individual characteristics of the first generation immigrants were included as explanatory variable of the perceived discrimination. In addition to the individual characteristics, the subsequent models separately include country-level ATII indicators (columns 2-7), the MIPEX scores (column 8), and the ATII indicators in combination with the MIPEX scores (columns 9-14).

[35] The native population and the second generation immigrants were excluded from regression analyses because of the inapplicability of some control variables (for example, the years of residency in the host country or the origin-to-host country controls) for these groups.

Table 3.4: OLS Regression Results of the Structure of Perceived Discrimination.

Perceived Discrimination	ATII Indicators							MIPEX Indicators						
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Young (< 25 y-o)	0.0546**	0.0550**	0.0548**	0.0554**	0.0544**	0.0545**	0.0545**	0.0548**	0.0556**	0.0553**	0.0556**	0.0549**	0.0547**	0.0547**
	(2.67)	(2.69)	(2.68)	(2.71)	(2.66)	(2.66)	(2.66)	(2.68)	(2.72)	(2.70)	(2.72)	(2.69)	(2.67)	(2.68)
Years of Residence	-0.000242	-0.000225	-0.000225	-0.000229	-0.000238	-0.000242	-0.000237	-0.000193	-0.000188	-0.000184	-0.000192	-0.000197	-0.000194	-0.000193
	(-0.64)	(-0.59)	(-0.59)	(-0.60)	(-0.62)	(-0.64)	(-0.62)	(-0.51)	(-0.49)	(-0.48)	(-0.50)	(-0.52)	(-0.51)	(-0.51)
Education (years)	0.00261*	0.00262*	0.00264*	0.00265*	0.00262*	0.00262*	0.00263*	0.00267*	0.00263*	0.00265*	0.00265*	0.00267*	0.00267*	0.00267*
	(2.47)	(2.47)	(2.49)	(2.51)	(2.47)	(2.47)	(2.48)	(2.52)	(2.49)	(2.50)	(2.50)	(2.52)	(2.52)	(2.52)
Female	-0.00771	-0.00820	-0.00825	-0.00836	-0.00778	-0.00786	-0.00788	-0.00705	-0.00736	-0.00742	-0.00750	-0.00693	-0.00710	-0.00708
	(-0.89)	(-0.95)	(-0.96)	(-0.97)	(-0.90)	(-0.91)	(-0.91)	(-0.82)	(-0.85)	(-0.86)	(-0.87)	(-0.80)	(-0.82)	(-0.82)
Married	-0.00736	-0.00712	-0.00710	-0.00689	-0.00737	-0.00733	-0.00724	-0.00674	-0.00640	-0.00634	-0.00616	-0.00672	-0.00670	-0.00671
	(-0.78)	(-0.75)	(-0.75)	(-0.73)	(-0.78)	(-0.77)	(-0.76)	(-0.71)	(-0.68)	(-0.67)	(-0.65)	(-0.71)	(-0.71)	(-0.71)
Presence of Children	-0.0175*	-0.0164	-0.0166*	-0.0168*	-0.0173*	-0.0174*	-0.0174*	-0.0168*	-0.0163	-0.0163	-0.0165	-0.0169*	-0.0168*	-0.0168*
	(-1.74)	(-1.63)	(-1.66)	(-1.67)	(-1.73)	(-1.73)	(-1.73)	(-1.67)	(-1.63)	(-1.62)	(-1.64)	(-1.69)	(-1.67)	(-1.67)
Naturalized	-0.0353***	-0.0348**	-0.0353***	-0.0350**	-0.0354***	-0.0354***	-0.035***	-0.0348**	-0.0342**	-0.0345**	-0.0341**	-0.0346**	-0.0348**	-0.0349**
	(-3.31)	(-3.26)	(-3.32)	(-3.29)	(-3.32)	(-3.32)	(-3.33)	(-3.27)	(-3.21)	(-3.24)	(-3.20)	(-3.25)	(-3.27)	(-3.27)
ATII INDICATORS														
Willingness to Allow More Immigrants														
of the same race		-0.202**							-0.178*					
		(-3.08)							(-2.20)					
of different race			-0.169*							-0.191*				
			(-2.40)							(-2.02)				
from poorer countries				-0.207**							-0.229*			
				(-3.02)							(-2.55)			
Opinion about Contribution to														
Economy					-0.0175							0.0305		
					(-0.72)							(1.02)		
Culture						-0.0219							-0.0118	
						(-0.62)							(-0.25)	
Country							-0.0378							-0.00930
							(-1.04)							(-0.21)
MIPEX SCORE														
Eligibility								-0.00384	-0.00559*	-0.00537*	-0.00510	-0.00427	-0.00391	-0.00378
								(-1.23)	(-1.73)	(-1.67)	(-1.61)	(-1.35)	(-1.24)	(-1.20)
Conditions for Acquisition								0.00305	0.00157	0.00320	0.00428	0.00302	0.00319	0.00323
								(0.82)	(0.41)	(0.86)	(1.14)	(0.81)	(0.85)	(0.84)
Security of Status								-0.0130*	-0.00784	-0.00973	-0.0106*	-0.0165*	-0.0123*	-0.0126*
								(-2.09)	(-1.18)	(-1.51)	(-1.69)	(-2.32)	(-1.79)	(-1.93)
Dual Nationality								0.0159**	0.00973	0.0105*	0.0102*	0.0189**	0.0152*	0.0155*
								(2.83)	(1.55)	(1.68)	(1.68)	(2.98)	(2.42)	(2.55)
Model Includes														
ATII Indicators	No	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
MIPEX Scores	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8539	8539	8539	8539	8539	8539	8539	8539	8539	8539	8539	8539	8539	8539
Adjusted R-squared	0.074	0.075	0.075	0.075	0.074	0.074	0.074	0.076	0.076	0.076	0.077	0.076	0.076	0.076

T statistics in parentheses

*** p<0.001, ** p<0.01, * p<0.1

Note: The outcome variable in all regression models is a binomial variable responsible for the perceived discrimination of the immigrants. The ATII indicators are country-average values of the responses of the native population computed separately for each ESS-round, and are limited to the native populations' average willingness to allow more immigrants of the same race as the majority, of a different race from the majority and from poorer countries outside Europe, as well as the native populations' average opinion about immigrants' contribution to the economy, culture and the host country in general. The MIPEX indicators are the values of the respective MIPEX scores of the "Eligibility", "Conditions for Acquisition", "Security of Status" and "Dual Nationality" for each year and country. The models with the MIPEX index scores also control for changes in the score values within 4 years prior to the survey. All regression models also include controls for origin to destination country pairs, and the data waves.

The results suggest that younger and more educated first generation immigrants experience higher perceived discrimination in the host countries. According to my expectations and the predictions of previous literature, naturalized immigrants experience lower perceived discrimination than non-naturalized immigrants. However, one cannot exclude the possibility that it is the socially more integrated immigrants that decide to naturalize, and that this finding may be a result of selection to naturalization rather than the naturalization itself. Nevertheless, the results suggest that naturalized immigrants are less likely to experience perceived discrimination compared to non-naturalized immigrants.

Surprisingly, the results show that the duration of residency in the host countries does not have any significant impact on the perceived discrimination of immigrants in those countries. This finding contradicts the assumption of better social acceptance of immigrants over time by the native population and stands robust when including more sets of controls (columns 2-14).

The regression coefficients of the ATII and MIPEX indicators also reveal some interesting and generally consistent results between the regression models. However, the table shows that the sign of the relationship may vary depending on the actual indicator that is being used.

In models with only ATII indicators (columns 2-7), the respective coefficients of all three indicators of the natives' willingness to allow more immigrants into their country show negative signs. This finding suggests that immigrants feel less discriminated in societies that are more willing to accept immigrants. Moreover, the coefficients keep their negative signs and statistical significance in models when controls for the stringency of the naturalization policies are introduced (columns 9-14). However, the indicators of the natives' average opinion about the contributions of immigrants to different aspects of the country did not reveal any statistically significant relationship with the perceived discrimination of immigrants.

Finally, the regression coefficients of the MIPEX scores of "Access to Nationality" (columns 8-14) show that the "Security of Status" is significantly and negatively correlated with the perceived discrimination of immigrants, which is consistent between models. This finding suggests lower perceived discrimination in countries where the naturalization status of immigrants is more secured. On the other hand, the positive coefficients of the indicator of "Dual Nationality", which was also statistically significant and mostly consistent between the models, suggest that immigrants experience more discrimination in countries where they are allowed to keep their initial nationality.

The results from the models in columns 9 and 10 also show negative relationships between the "Eligibility" score and the perceived discrimination of immigrants. This suggests lower perceived discrimination of immigrants in countries with more favorable eligibility conditions for naturalization of immigrants. However, similar to the coefficients of the indicator of "Conditions for Acquisitions", these results lose their statistical significance in other models. Nevertheless, the general results of this sub-section are in line with the hypothesis that the attitude of the native population towards immigrants and the favorability of naturalization policies for immigrants are negatively correlated with the perceived discrimination of first generation immigrants in the host courtiers.

3.5.3 Social Integration of Immigrants

This sub-section covers the indicators of the social integration of immigrants expressed through the feeling of closeness to the country of residence and a measure of the frequency by which the official language of the host country is being spoken at home. The causality and the direction of the relationships of these indicators with the ATII indicators and the MIPEX scores are again confusing, similar to those in the previous sub-section. Moreover, the impact of the stringency of naturalization policies on the social integration of immigrants could still simply capture the indirect impact of ATII.

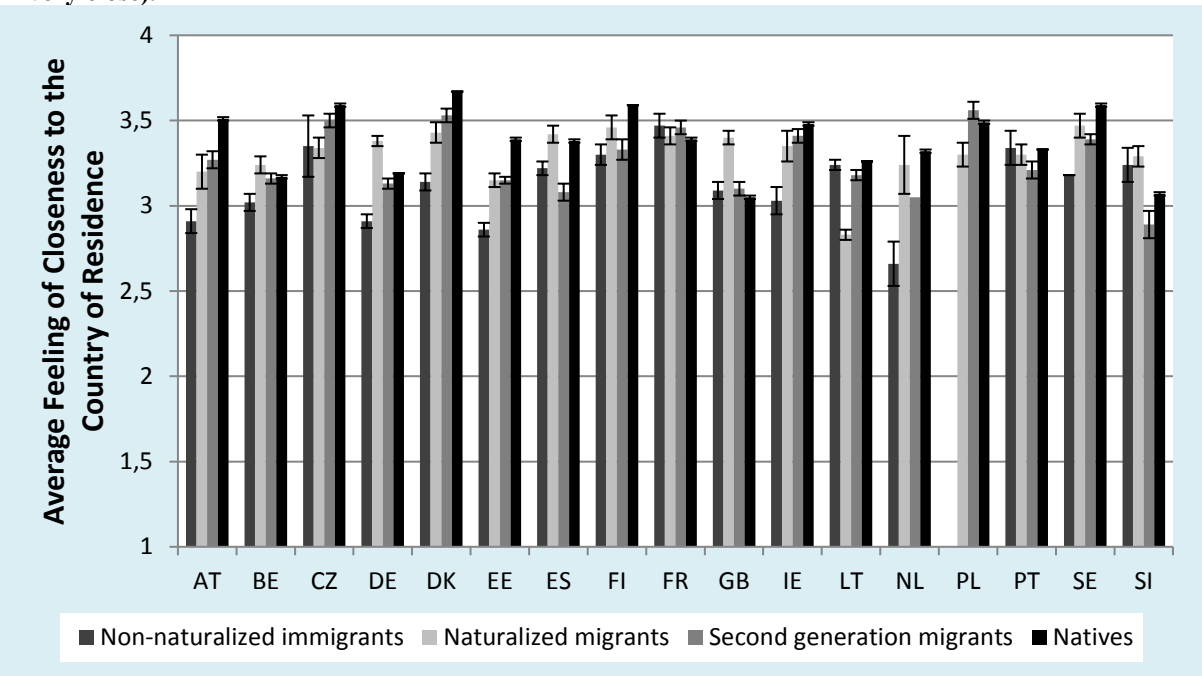
Figure 3.3.1 presents the group-average subjective feeling of closeness to the country of residence[36]. This indicator was constructed by averaging the individuals' answers to the question as to how close they feel to the country of residence, with four possible answers ranging from 1 for "not close at all" to 4 for "very close".

The results do not show much economically significant within-country variations in the feeling of closeness to the country between the native population and different groups of immigrants. However, statistically significant differences are noticeable for most countries, particularly between the groups of natives and non-naturalized immigrants. Very interesting results are obtained for France, Sweden and the UK, where the average reported feeling of closeness of the native population is less than those of some groups of immigrants. This issue

[36] The results of GR, IT and SK are absent, since this question was added to the ESS questioner only for the 2014 data wave, and the corresponding data wave for these countries is missing.

raises some concerns regarding the attitude of the native population towards their country, particularly in the UK where the feeling of closeness of the native population is the lowest among the countries in the figure.

Figure 3.3.1: The Group-Average "Feeling of Closeness" to the Country of Residence (1=not close at all, 4=very close).



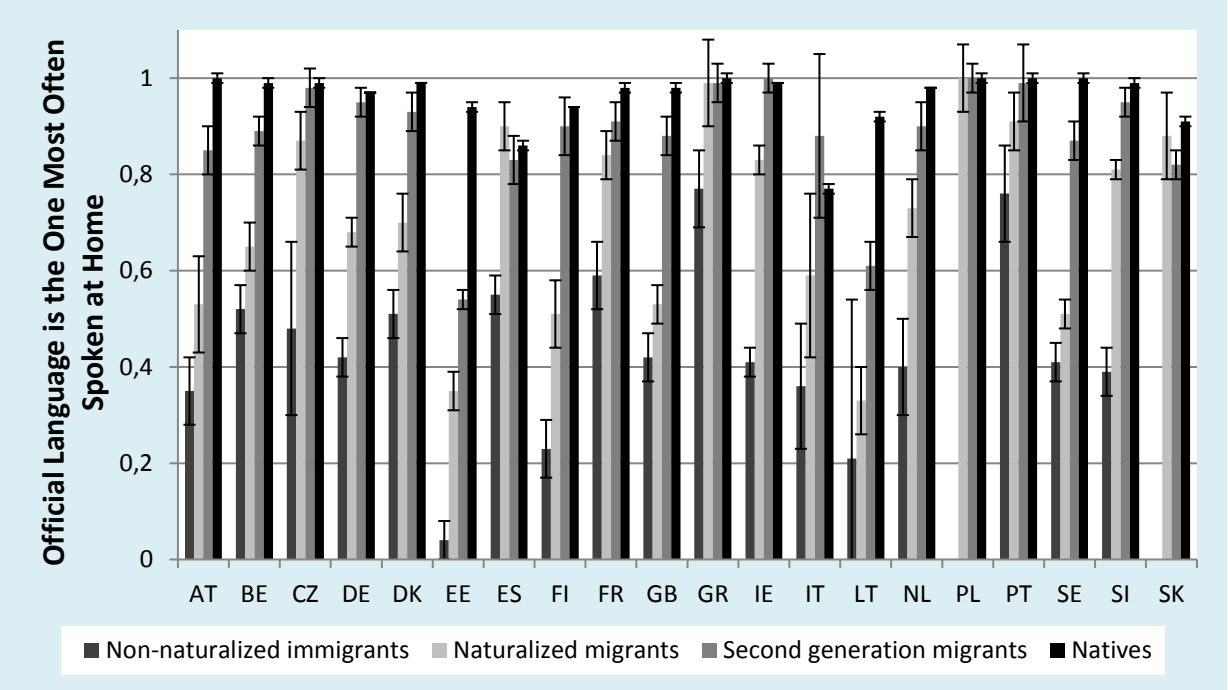
The upper and lower bars stand for a 95% confidence interval.

Note: The survey question used for the creation of this indicator was: "How close do you feel to the country?" The scaling of this survey question could take integer values ranging from 1 for "not close at all" to 4 for "very close". Only the 2014-year wave contains this question and thus GR and IT estimates are missing from the figure. The group averages of this indicator were separately computed for the groups of non-naturalized immigrants, naturalized immigrants, second generation immigrants, and the native population.

Figure 3.3.2 presents per-country levels of the likelihood that the official language of the host country is that which is most often spoken at home. Language acquisition is one of the most important indicators of acculturation used in the literature and is highly appropriate in the evaluation of the social integration of immigrants.

The figure shows that the countries with the most unfavorable citizenship policies for immigrants (AT, EE and LT) have some of the lowest rates of language acquisition both for naturalized and second generation immigrants. The countries with the highest levels of language acquisition of naturalized immigrants are CZ, ES, FR, GR, PL and PT which, based on the composition of immigrants in these countries, may be a result of the closeness of the host and origin country languages rather than the acculturation itself.

Figure 3.3.2: The Group-Average Probability that the Official Language of the Country of Residence is that which is Most Often Spoken at Home.



The upper and lower bars stand for a 95% confidence interval.

Note: This indicator was constructed using the survey question "What is the language most often spoken at your household?", and matching the answers with the official language or languages of the destination countries. The scaling of this indicator could take values "1" (if the official language of the country was the one most often spoken at home) and "0" otherwise. The group averages of this indicator were separately computed for the groups of non-naturalized immigrants, naturalized immigrants, second generation immigrants, and the native population.

Similar to the previous sub-section, regression analyses were conducted on the joint sample of first generation naturalized and non-naturalized immigrants in order to reveal the relationships between the social integration of immigrants and the ATII indicators or the MIPEX sores of the stringency of naturalization policies. Since it is impossible to distinguish the causality of these relationships with the ESS data, I will again focus only on the directions of the relationships between the chosen indicators.

Unfortunately, since the question regarding the feeling of closeness to the country of residence was added to the ESS questionnaire only in 2014, the respective regression analyses were not carried out for this indicator because of the low number of observations. However, Table 3.5 presents the results of OLS regression analyses for the likelihood that the official language of the host country is that which is most often spoken at home.

As opposed to the results regarding the perceived discrimination of immigrants in the previous sub-section, the results in Table 3.5 show that alongside naturalization, the duration of residence in the host countries has a statistically significant positive relationship with language acquisition.

Table 3.5: OLS Regression Results of Language Acquisition.

Language Acquisition	ATII Indicators							MIPEX Indicators						
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Young (< 25 y-o)	0.0769***	0.0766***	0.0769***	0.0768***	0.0767***	0.0769***	0.0769***	0.0783***	0.0781***	0.0787***	0.0785***	0.0779***	0.0782***	0.0782***
	(3.62)	(3.61)	(3.62)	(3.62)	(3.61)	(3.62)	(3.62)	(3.69)	(3.68)	(3.71)	(3.70)	(3.67)	(3.69)	(3.68)
Years of Residence	0.00555***	0.00554***	0.00555***	0.00555***	0.00556***	0.00555***	0.00555***	0.00556***	0.00555***	0.00556***	0.00556***	0.00557***	0.00555***	0.00556***
	(14.02)	(14.00)	(14.02)	(14.02)	(14.03)	(14.02)	(14.02)	(14.03)	(14.02)	(14.04)	(14.03)	(14.06)	(14.02)	(14.04)
Education (years)	0.00475***	0.00474***	0.00474***	0.00474***	0.00475***	0.00474***	0.00475***	0.00467***	0.00469***	0.00466***	0.00467***	0.00467***	0.00465***	0.00467***
	(4.33)	(4.33)	(4.33)	(4.33)	(4.33)	(4.33)	(4.33)	(4.26)	(4.27)	(4.25)	(4.26)	(4.26)	(4.24)	(4.26)
Female	0.0387***	0.0389***	0.0387***	0.0387***	0.0386***	0.0387***	0.0387***	0.0388***	0.0389***	0.0387***	0.0388***	0.0385***	0.0386***	0.0387***
	(4.32)	(4.35)	(4.32)	(4.33)	(4.32)	(4.33)	(4.32)	(4.34)	(4.35)	(4.32)	(4.33)	(4.31)	(4.32)	(4.32)
Married	-0.0446***	-0.0448***	-0.0446***	-0.0447***	-0.0446***	-0.0447***	-0.0446***	-0.0447***	-0.0448***	-0.0445***	-0.0446***	-0.0447***	-0.0445***	-0.0445***
	(-4.56)	(-4.57)	(-4.56)	(-4.56)	(-4.56)	(-4.56)	(-4.56)	(-4.56)	(-4.57)	(-4.54)	(-4.55)	(-4.57)	(-4.53)	(-4.54)
Presence of Children	-0.00888	-0.00941	-0.00890	-0.00891	-0.00877	-0.00892	-0.00888	-0.00968	-0.00982	-0.00939	-0.00961	-0.00935	-0.00959	-0.00961
	(-0.86)	(-0.91)	(-0.86)	(-0.86)	(-0.85)	(-0.86)	(-0.86)	(-0.93)	(-0.95)	(-0.91)	(-0.93)	(-0.90)	(-0.93)	(-0.93)
Naturalized	0.127***	0.127***	0.127***	0.127***	0.127***	0.127***	0.127***	0.127***	0.127***	0.127***	0.127***	0.126***	0.127***	0.126***
	(11.52)	(11.50)	(11.52)	(11.52)	(11.52)	(11.53)	(11.52)	(11.48)	(11.46)	(11.49)	(11.49)	(11.44)	(11.47)	(11.45)
ATII INDICATORS														
Willingness to Allow More Immigrants														
of the same race		0.0993							0.0535					
		(1.47)							(0.64)					
of different race			0.00516							-0.103				
			(0.07)							(-1.05)				
from poorer countries				0.0102							-0.0425			
				(0.14)							(-0.46)			
Opinion about Contribution to														
Economy					-0.0128							-0.0772*		
					(-0.51)							(-2.50)		
Culture						0.00973							-0.0697	
						(0.27)							(-1.44)	
Country							0.00145							-0.0662
							(0.04)							(-1.41)
MIPEX SCORE														
Eligibility								0.00167	0.00220	0.000848	0.00143	0.00281	0.00126	0.00215
								(0.51)	(0.65)	(0.25)	(0.43)	(0.85)	(0.39)	(0.65)
Conditions for Acquisition								-0.00377	-0.00332	-0.00369	-0.00354	-0.00362	-0.00289	-0.00240
								(-0.97)	(-0.84)	(-0.95)	(-0.91)	(-0.93)	(-0.74)	(-0.60)
Security of Status								0.0118*	0.0102	0.0135*	0.0122*	0.0204**	0.0161*	0.0146*
								(1.82)	(1.48)	(2.03)	(1.87)	(2.79)	(2.26)	(2.16)
Dual Nationality								-0.00877	-0.00691	-0.0117*	-0.00983	-0.0163*	-0.0129*	-0.0120*
								(-1.51)	(-1.06)	(-1.82)	(-1.57)	(-2.49)	(-1.99)	(-1.92)
Model Includes														
ATII Indicators	No	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
MIPEX Scores	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8719	8719	8719	8719	8719	8719	8719	8719	8719	8719	8719	8719	8719	8719
Adjusted R-squared	0.418	0.418	0.418	0.418	0.418	0.418	0.418	0.419	0.419	0.419	0.419	0.419	0.419	0.419

T statistics in parentheses

*** p<0.001, ** p<0.01, * p<0.1

Note: The outcome variable in all regression models is a binomial variable responsible for the likelihood that the official language of the host country is the one that is most often spoken at home by immigrants. The ATII indicators are country-average values of the responses of the native population computed separately for each ESS-round, and are limited to the native populations' average willingness to allow more immigrants of the same race as the majority, of a different race from the majority and from poorer countries outside Europe, as well as the native populations' average opinion about immigrants' contribution to the economy, culture and the host country in general. The MIPEX indicators are the values of the respective MIPEX scores of the "Eligibility", "Conditions for Acquisition", "Security of Status" and "Dual Nationality" for each year. The models with the MIPEX index scores also control for changes in the score values within 4 years prior to the survey.
All regression models also include controls for origin to destination country pairs, and the data waves.

Other individual characteristics also revealed expected relationships with language acquisition. In particular, the results suggest that younger, higher educated and female immigrants are more likely to acquire the official language of the host country. On the other hand, married immigrants have significantly lower chances of acquisition of the host country language, while the presence of children did not reveal any statistically significant relationship.

Surprisingly, none of the regression coefficients of the ATII indicators showed any statistically significant relationship with the language acquisition. This result suggests that the attitude of the native population of a country does not relate to the acculturation of immigrants in that country.

Finally, the indicators of the MIPEX score of "Access to Nationality" revealed results that are consistent with the findings in the previous sub-section. In particular, in most models the score of "Security of Status" shows a statistically significant positive relationship with the language acquisition of immigrants, while the indicators of "Dual Nationality" show negative relationships. This finding is in line with the results in the previous sub-section that suggest better social integration of immigrants in countries where the naturalization status of immigrants is more secured but possibilities for dual nationality are more restricted.

The general findings of this sub-section again confirm that the relationships between ATII, naturalization policies, and the social integration of immigrants present a complex system. The ATII indicators did not reveal any statistically significant relationship with the acculturation of immigrants, whereas the MIPEX indicator of "Security of Status" revealed a positive relationship, while the indicator of "Dual Nationality" revealed a negative relationship with the social integration of immigrants.

3.6 Conclusion

This study contributes to the existing academic literature by being the first to analyze the perceived discrimination and social integration of first and second generation immigrants in combination with the attitude of the native population and institutionally imposed restrictions on naturalization.

The findings from the evaluation of the relationship between naturalization policies and the attitude of a country's native population towards immigrants and immigration suggest that the latter is fairly represented in a country's naturalization policies. In particular, the favorability

of a country's naturalization policies towards immigrants is closely and positively related to the natives' desire to allow more immigrants from poorer countries, the natives' desired level of government generosity in judging refugee applications, as well as the natives' opinion of immigrants' positive contribution to various aspects of the country. This finding in general supports the hypothesis of a positive correlation between ATII and the favorability of naturalization policies for immigrants, which is also suggested by the related literature (Facchini and Mayda, 2008; Facchini and Mayda, 2009; Facchini and Mayda, 2012).

As opposed to the logical expectation and the predictions in previous literature regarding the determinants of immigrant integration (Vink, Prokic-Breuer and Dronkers, 2013; Gathmann and Keller, 2014; Sargsyan, 2017), the results of the regression analyses suggest that the duration of residence in host countries is not related to a decline in the perceived discrimination of immigrants. However, in line with the predictions of the related literature, the duration of residence in host countries and naturalization of immigrants were shown to have statistically significant positive relationships with the chosen indicators of acculturation. Some other individual characteristics also revealed expected relationships with language acquisition but showed no significant impact on the perceived discrimination of immigrants.

In further evaluation of the perceived discrimination of immigrants, the regression results revealed that it is negatively related to the natives' average willingness to allow more immigrants of different origins into their country. This finding suggests that immigrants feel less discriminated against in societies that are more welcoming. However, despite being in line with the predictions and hypothesis in the related literature (Vink, Prokic-Breuer and Dronkers, 2013; Gathmann and Keller, 2014; Sargsyan, 2017), it is difficult to identify the causality of this negative relationship.

On the other hand, indicators of the stringency of naturalization policies revealed both positive and negative relationships with the perceived discrimination of immigrants depending on the actual indicator used. In particular, the MIPEX indicator of "Security of Status" revealed a significant negative relationship with the perceived discrimination of immigrants, while the indicator of "Dual Nationality" revealed a positive relationship. This finding suggests lower perceived discrimination of immigrants in countries where the naturalization status of immigrants is more secured, and the possibility of dual nationality is more restricted. Moreover, language acquisition of immigrants was also positively related with the security of the

naturalization status and negatively related with the indicator of dual nationality. However, the ATII indicators did not show any statistically significant relationship with language acquisition.

The general findings of this study again confirm that the relationships between ATII, naturalization policies, and the social integration of immigrants present a complex system with a wide range of interconnected mechanisms. More research is needed to reveal the exact mechanisms of these relationships.

3.A Appendix

Table 3.6: The Scores of the Sub-categories of the MIPEX Score of "Access to Nationality".

	MIPEX Score of "Eligibility"					MIPEX Score of "Conditions for Acquisition"				
ISO	2010	2011	2012	2013	2014	2010	2011	2012	2013	2014
AT	21	21	21	21	21	18	17	17	17	17
BE	67	67	67	63	63	67	67	67	28	28
CZ	0	0	0	0	33	62	62	62	62	28
DE	92	92	92	92	92	52	52	52	52	52
DK	46	46	46	63	63	30	32	32	42	42
EE	8	8	8	8	8	42	42	42	42	42
ES	50	50	50	50	50	32	32	32	32	32
FI	71	79	79	79	79	58	58	58	58	58
FR	79	79	79	79	79	23	23	20	25	25
GB	79	79	79	79	79	35	35	35	27	27
GR	71	71	71	29	29	37	37	37	37	37
IE	88	96	96	96	96	33	33	33	33	33
IT	42	42	42	42	42	23	23	23	23	23
LT	25	25	25	25	25	58	58	58	58	58
NL	71	71	71	71	71	48	48	48	42	42
PL	10	10	20	20	20	62	62	63	63	63
PT	92	92	92	92	92	80	80	80	80	78
SE	50	50	50	50	50	83	83	83	83	83
SI	21	21	21	21	21	53	53	53	53	53
SK	8	8	8	8	8	15	15	15	15	15

	MIPEX Score of "Security of Status"					MIPEX Score of "Dual Nationality"				
ISO	2010	2011	2012	2013	2014	2010	2011	2012	2013	2014
AT	30	30	30	30	30	38	38	38	38	38
BE	13	13	13	83	83	100	100	100	100	100
CZ	60	60	60	60	60	38	38	38	38	75
DE	57	57	57	57	57	63	63	63	63	88
DK	27	27	27	27	27	38	38	38	38	100
EE	20	20	20	20	20	0	0	0	0	0
ES	60	60	60	60	60	50	50	50	50	50
FI	40	40	40	40	40	75	75	75	75	75
FR	40	40	40	40	40	100	100	100	100	100
GB	33	33	33	33	33	100	100	100	100	100
GR	20	20	20	20	20	100	100	100	50	50
IE	7	7	7	7	7	100	100	100	100	100
IT	60	60	60	60	60	75	75	75	75	75
LT	10	30	30	30	30	25	25	25	25	25
NL	63	63	63	63	63	88	88	88	88	88
PL	20	20	90	90	90	13	13	50	50	50
PT	73	73	73	73	73	100	100	100	100	100
SE	60	60	60	60	60	100	100	100	100	100
SI	53	53	53	53	53	38	38	38	38	38
SK	27	27	27	27	27	50	50	50	50	50

Table 3.7: The Sub-categories of the MIPEX Score of "Access to Nationality".

Sub-category	Covered Question	Included Indicators
Eligibility	How long must migrants wait to naturalize? Are their children and grandchildren born in the country entitled to become citizens?	Residence period; Permits considered; Periods of prior-absence allowed; Requirements for spouses; Requirements for partners; Birth-right citizenship for second generation.
Conditions for Acquisition of Status	Are applicants encouraged to succeed through basic conditions for naturalization?	Naturalization language requirement; Integration requirement (form, exemption, cost, support, courses); Economic resources; Criminal record; Good character; Cost of application.
Security of Status	Does the state protect applicants from discretionary procedures?	Maximum duration of procedure; Additional grounds for refusal; Discretionary powers in refusal; Legal protection; Protection against withdrawal of citizenship.
Dual Nationality	Can naturalizing migrants and their children be citizens of more than one country?	Dual nationality for first generation (Renunciation requirement, Renunciation exemptions); Dual nationality for second generation; Dual nationality for third generation.

Table 3.8: Correlation Table between the ATII Indicators.

		Willingness to Allow More Immigrants			Natives' Opinion about Contribution to		
		of the same race	of different race	from poorer countries	Economy	Culture	Country
Willingness to Allow More Immigrants	of the same race	1.00					
	of different race	0.88	1.00				
	from poorer countries	0.78	0.94	1.00			
Natives' Opinion about Contribution to	Economy	0.74	0.82	0.72	1.00		
	Culture	0.73	0.86	0.80	0.86	1.00	
	Country	0.71	0.85	0.74	0.83	0.90	1.00

Summary

In this study, I provide a more comprehensive picture of the integration of immigrants in host societies and labor markets by first studying the importance of governmental registration in the successful socio-economic integration of immigrants in destination societies, and then attempting to link the attitude of the native population and institutionally-imposed restrictions in the form of naturalization policies with the social integration of immigrants.

In the first chapter, I introduce a new dimension for income decomposition in China, through testing whether there is differential treatment towards rural-urban migrants in urban China who converted the type of their official registration (*hukou*) to local-urban registration. I also test how the treatment towards *hukou* converters differs from the treatment towards those who did not convert their registration type.

The findings of the Oaxaca-Blinder decomposition analysis indicate that, despite its huge assumed impact on the differential treatment towards rural migrant workers in urban China, the type of *hukou* is not entirely responsible for the unexplainable disadvantageous treatment towards migrants working in the paid-employment sector. Measurable "unexplained" components of wage gaps were found for the Converted-Rural couples for workers of both genders, and particularly for migrant males working in the paid-employment sector. This finding suggests that reforms in the *hukou* registration system, or even the complete abolishment of this system, may not actually cause the elimination of the differential treatment towards rural migrants in Chinese labor markets.

In the second chapter, I further extend the analyses of the integration of migrants in labor markets by focusing on international immigrants in two European countries with different naturalization policies (France and Denmark). In particular, I conduct a wage decomposition analysis for naturalized and non-naturalized immigrants and the native population at destination and compare the returns to the average personal characteristics of individuals belonging to each of the three groups in the two countries.

The "unexplained" components in the wage differences between the groups suggest the existence of high naturalization premiums in France, but do not show any naturalization premiums in Denmark. The results also suggest that after naturalization migrants become highly integrated in the French labor market and are treated similarly to natives. However, this is not the case in Denmark. Since Denmark has stricter naturalization policies compared to

France, these findings question the assumption that stricter naturalization policies result in better socioeconomic integration of migrants. Indeed, the findings suggest that it may be the socio-cultural attitude to immigrants that results in the formation and strictness level of naturalization policies.

Finally, in the third chapter, I attempt to test the above hypotheses, by analyzing the relationships between the attitude of the native population towards immigrants in European countries and "social integration" of first and second generation immigrants in those countries. The stringency of naturalization policies in the host countries is also taken into consideration as a mechanism of the development of this relationship.

The results show that a country's naturalization policies fairly represent some ATII indicators such as natives' willingness to allow more immigrants from poorer countries and natives' desired level of government generosity in judging refugee applications.

The findings from the evaluation of the relationship between naturalization policies and the attitude of a country's native population towards immigrants and immigration suggest that the latter is fairly represented in a country's naturalization policies. Moreover, the natives' willingness to accept new immigrates, which were among the chosen ATII indicators, showed significant negative relationships with the perceived discrimination of immigrants, indicating that immigrants feel less discriminated against in more welcoming societies. However, the ATII indicators did not reveal significant consistent relationships with the indicators of the social integration of immigrants. On the other hand, the findings from the MIPEX indicator suggest lower perceived discrimination and higher social integration of immigrants in countries where the naturalization status of immigrants is more secured, and the possibility of dual nationality is more restricted.

The duration of residence in the host countries did not reveal a significant relationship with the perceived discrimination of immigrants, but it showed statistically significant positive relationships with the chosen indicators of acculturation.

In general, the findings of the dissertation again confirm that immigrant integration is a complex process with a wide range of interconnected mechanisms that shape its relationship with the attitude of the native population or institutionally-imposed restrictions.

References

Akay, A., Bargain, O., and Zeimmermann, K. (2012). "Relative Concerns of Rural-to-Urban Migrants in China". *Journal of Economic Behavior and Organization*, *81, 421-441.*

Akgüç, M., Giulietti, C., and Zimmermannal, K. F. (2014). "The RUMiC longitudinal survey: fostering research on labor markets in China". *IZA Journal of Labor & Development*, *3:5.*

Aleksynska, M., and Algan, Y. (2010). "Assimilation and Integration of Immigrants in Europe". IZA Discussion Paper no. 5185.

Algan, Y., Dustmann, C., Glitz, A., and Manning, A. (2010). "The Economic Situation of First and Second-Generation Immigrants in France, Germany and The United Kingdom". *The Economic Journal, 120: F4-F30.*

Arrow, K. (1973). "The Theory of Discrimination". Discrimination in Labor Markets, *Princeton University Press, 3-33.*

Beblo, M., Beninger, D., Heinze, A., and Laisney, F. (2003). "Methodological Issues Related to the Analysis of Gender Gaps in Employment, Earnings and Career Progression". European Commission Report.

Bernitz, H. (2012). "Research for the EUDO Citizenship Observatory. Country Report: Sweden". European University Institute.

Bertossi, C. (2010). "Research for the EUDO Citizenship Observatory Country Reports. Report on France". European University Institute.

Bertossi, C., and Hajjat, A. (2013). "Research for the EUDO Citizenship Observatory. Country Report: France". European University Institute.

Bevelander, P., and Pendakur, R. (2012). "Citizenship Acquisition, Employment Prospects and Earnings: Comparing Two Cool Countries". Working paper, Robert Schuman Centre, Florence.

Bevelander, P., and Veenman, J. (2008). "Naturalization and Socioeconomic Integration: The Case of the Netherlands". IZA Discussion Paper No. 2153.

Blinder, A. S. (1973). "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources* 8*, 436-455.*

Bond, T. N., and Lang, K. (2013). "The Evolution of the Black-White Test Score Gap in Grades K–3: The Fragility of Results". *The Review of Economics and Statistics 95(5), p. 1468–1479.*

Bohman, A., and Hjerm, M. (2016). "In the Wake of Radical Right Electoral Success: A Cross-Country Comparative Study of Anti-Immigration Attitudes Over Time". *Journal Of Ethnic And Migration Studies, 42/ 11, 1729–174.*

Borjas, G. J. 2008. "Labor Outflows and Labor Inflows in Puerto Rico". *Journal of Human Capital 2:32–68.*

Bratsberg, B., Ragan, J.F., and Nasir, Z.M. (2002). "The Effect of Naturalization on Wage Growth: A Panel Study of Young Male Immigrants". *Journal of Labor Economics 20, no. 3:568-597.*

Callens, M.-S. (2014). “Do Attitudes Towards the Integration of Immigrants Change Over Time? A Comparative Study of Natives, Second-Generation Immigrants and Foreign-Born Residents in Luxembourg”. *Journal of International Migration & Integration 15/1: p135-157.*

Casarico, A., Facchiniy, G., and Frattini, T. (2015). “Illegal Immigration: Policy Perspectives and Challenges”. *CESifo Economic Studies Advance Access 61(3-4), p. 673–700.*

Ceobanu, A. M., and Escandell, X., (2005). “East is West? National feelings and anti-immigrant sentiment in Europe”. *Social Science Research 37: 1147-1170*.

Ceobanu, A. M., and Escandell, X., (2010). “Comparative Analyses of Public Attitudes Toward Immigrants and Immigration Using Multinational Survey Data: A Review of Theories and Research”. *Annual Review of Sociology 36:1, 309-328.*

Contucci, P., and Sandell, R. (2015). “How Integrated are Immigrants?” *Demographic Research 33/46: 1271-1280*.

Cheng, Z., Guo, F., Hugo, G., and Yuan X. (2013). “Employment and Wage Discrimination in the Chinese Cities: A Comparative Study of Migrants and Locals”. *Habitat International, 39, 246-255.*

Christofides, L. N., Polycarpo, A., and Vrachimis, K. (2013). “Gender Wage Gaps, ‘Sticky Floors’ and ‘Glass Ceilings’ in Europe”. *Labor Economics, 21, 86-108.*

Démurger, S., Gurgand, M., Li, S., and Ximing, Y. (2009). “Migrants as Second-class Workers in Urban China? A Decomposition Analysis”. *Journal of Comparative Economics*, 37*, 610-628.*

Démurger, S., and Xu, H. (2013). “Left-Behind Children and Return Decisions of Rural Migrants in China”. IZA discussion paper no. 7727.

Dong, X.Y., and Bowles, P. (2002). “Segmentation of Discrimination in China’s Emerging Industrial Labor Market”. *China Economic Review, 13, 170-196.*

Dustmann, C., and Frattini, T. (2011). “Immigration: The European Experience”. Centre for Research and Analysis of Migration (CReAM) Discussion Paper No 22/11.

Dustmann, C., Glits, A., and Voge, T. (2009). “Employment, Wages, and the Economic Cycle: Differences between Immigrants and Natives”. IZA Discussion Paper No. 4432.

DeVoretz, D. J., and Pivnenko, S. (2004). “The Economic Causes and Consequences of Canadian Citizenship”. *Journal of International Migration and Integration 6:435–468.*

Dronkers, J., and Vink, M.P. (2012). “Explaining Access to Citizenship in Europe: How Citizenship Policies Affect Naturalization Rates“. *European Union Politics 13, no. 3:390-412*.

Elsuwege, V. P. (2004). “Russian-Speaking Minorities in Estonia and Latvia: Problems of Integration at the Threshold of the European Union”. ECMI Working Paper No. 20.

Engdahl, M. (2011). “The Impact of Naturalization on Labor Market Outcomes in Sweden, in OECD, Naturalization: A Passport for the Better Integration of Immigrants?” OECD Publishing.

Entzinger, H., and Biezeveld R. (2003). “Benchmarking in Immigrant Integration”. European Research Centre on Migration and Ethnic Relations (ERCOMER).

Ersboll, E. (2010). “Research for the EUDO Citizenship Observatory Country Reports”. Report on Denmark. European University Institute.

Facchini, G., and Mayda, A. M. (2008). “From Individual Attitudes Towards Migrants to Migration Policy Outcomes: Theory and Evidence”. *Economic Policy 56:651–713.*

Facchini, G., and Mayda, A. M. (2009). “Does the Welfare State Affect Individual Attitudes Toward Immigrants? Evidence Across Countries”. *The Review of Economics and Statistics 91:295–314.*

Facchini, G., and Mayda, A. M. (2012). “Individual Attitudes Towards Skilled Migration: An Empirical Analysis Across Countries”. *The World Economy 35(2):183–96.*

Fagerlund, J., and Brander, S. (2013). “Research for the EUDO Citizenship Observatory. Country Report: Finland”. European University Institute.

Foblets, M. C., Yanasmayan, Z., and Wautelet P. (2013). “Research for the EUDO Citizenship Observatory. Country Report: Belgium”. European University Institute.

Fougere, D., and Safi, M. (2008). “the Effects of Naturalization on Immigrants’ Employment Probability (France, 1968 – 1999)”. IZA Discussion Paper No. 3372.

Gagnon, J., Xenogiani, T., and Xing, C. (2014). “Are All Migrants Really Worse Off in Urban Labor Markets? New Empirical Evidence from China”. IZA Discussion Paper No. 6268.

Gardner, P.L. (1975), “Scales and statistics”. *Renew of Educational Research, 45, 43-57.*

Gathmann, C., and Keller, N. (2014). “Returns to Citizzenship? Evidence from Germany’s Recent Immigration Reforms”. IZA Discussion Paper No. 8064.

Giulietti C., Ning G., Zimmermann K.F. (2012). "Self- employment of rural- to- urban migrants in China". *International Journal of Manpower, Vol. 33 Issue: 1, pp.96-117.*

Gorinas, C., and Pytliková M. (2017), “The Influence of Attitudes toward Immigrants on International Migration”. *International Migration Review, 51(2), 416–451.*

Gui, Y., Berry, J., W., and Zheng, Y. (2012). “Migrant Worker Acculturation in China”. *International Journal of Intercultural Relations*, 36, 598-610.

Guzi, M., and Kahanec, M. (2017). "How immigrants helped EU labor markets to adjust during the Great Recession", *International Journal of Manpower, 38(7): 996-1015*.

Guzi, M., Kahanec, M., and Kureková, L.M. (2015). "What Explains Immigrant-Native Gaps in European Labor Markets: The Role of Institutions". IZA Discussion Paper No. 8847.

Hailbronner, K., and Farahat, A. (2015). “Research for the EUDO Citizenship Observatory. Country Report on Citizenship Law: Germany”. European University Institute.

Hamilton, B.H. (2000). “Does entrepreneurship pay? An empirical analysis of returns to self-employment”. *Journal of Political Economy, 108, 604–632.*

Handoll, J. (2012). “Research for the EUDO Citizenship Observatory. Country Report: Ireland”. European University Institute.

Hatton, T. J. (2014). “Public Opinion on Immigration: Has the Recession Changed Minds?”. CEPR Discussion Paper No. DP10008. Available at SSRN: https://ssrn.com/abstract=2501460.

Hatton, T. J. (2017). “Public Opinion on Immigration in Europe: Preference Versus Salience”. CEPR Discussion Paper No. DP12084. Available at SSRN: https://ssrn.com/abstract=2984780.

Heckman, J.J. (1979). “Sample Selection Bias as a Specification Error”. *Econometrica, 47, 153-163.*

Helgertz, J., Bevelander, P., and Tegunimataka, A. (2014). “Naturalization and Earnings: A Denmark–Sweden Comparison”. *European Journal of Population 30:337–359.*

Huddleston, T., Niessen, J., Ni Chaoimh E., and White, E. (2011). “Migrant Integration Policy Index III”. Brussels: British Council and Migration Policy Group.

Jarve, P., and Poleschuk, V. (2013). “Research for the EUDO Citizenship Observatory. Country Report: Estonia”. European University Institute.

Keller, N., Gathmann, C., and Monscheuer, O. (2015). “Citizenship and the Social Integration of Immigrants: Evidence from Germany's Immigration Reforms”. Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Labor - Regulatory Effects, No. A18-V4

Kovacs, M. M., and Toth, J. (2010). “Research for the EUDO Citizenship Observatory Country Reports. Report on Hungary”. European University Institute.

Lee, L. (2012). “Decomposing Wage Differentials between Migrant Workers and Urban Workers in Urban China’s Labor Markets”. *China Economic Review, 23, 461-470.*

Lu, Z., and Song, S. (2006). “Rural-Urban Migration and Wage Determination: The Case of Tianjin, China”. *China Economic Review, 17, 337-345.*

Magnani, E., and Zhu, R. (2012). “Gender Wage Differentials among Rural-Urban Migrants in China”. *Regional Science and Urban Economics, 42, 779-793.*

Manning, A. (2010). “The Integration of Immigrants and Their Children in Europe: Introduction”. *The Economic Journal, 120/ 542: F1-F3.*

Marin, R. R., Sobrino, I., Perez, A. M. and Fuentes, F. J. M. (2015). “Research for the EUDO Citizenship Observatory. Country Report on Citizenship Law: Spain”. European University Institute.

Mayda, A. M. (2006). “Who Is Against Immigration? A Cross-Country Investigation of Individual Attitudes toward Immigrants”. *Review of Economics and Statistics 88/3: 510-530.*

Meng, X., and Zhang, J. (2001). “The Two-tier labor Market in Urban China: Occupational Segregation and Wage Differentials between Urban Residents and Rural Migrants in Shanghai”. *Journal of Comparative Economics, 29, 485-504.*

Meuleman, B., Davidov, E., and Billiet, J. (2009). “Changing Attitudes toward Immigration in Europe, 2002–2007: A Dynamic Group Conflict Theory Approach”. *Social Science Research 38:352-365.*

Mincer, J. (1958). “Investment in Human Capital and Personal Income Distribution”. *Journal of Political Economy, 66(4), 281-302.*

McKenzie, D., and Rapoport, H. (2010). “Self-selection patterns in Mexico-U.S. migration: the role of migration networks”. *The Review of Economics and Statistics 92:811–821.*

Neuman, S., and Oaxaca, R. L. (2001). “Estimating Labor Market Discrimination with Selectivity-Corrected Wage Equations: Methodological Considerations and an Illustration from Israel”. Mimeo.

Neumark, D. (1988). “Employers' Discriminatory Behavior and the Estimation of Wage Discrimination”. *Journal of Human Resources, 23, 379-295.*

Oaxaca, R. L. (1973). “Male-Female Wage Differentials in Urban Labor Markets,” *International Economic Review, 14, 693-709.*

Oaxaca, R., L., and Ransom, M., R. (1994). “On Discrimination and Decomposition of Wage Differentials”. *Journal of Econometrics*, *61, 5-21.*

Ortega, F., and Peri G. (2013). “The Effect of Income and Immigration Policies on International Migration”. *Migration Studies 1:1–28.*

Park A and Wang D (2010) “Migration and urban poverty and inequality in China”. *China Economic Journal, 3: 49–67.*

Pastor, M., and Scoggins, J. (2012). “The Economic Benefits of Naturalization for Immigrants and the Economy. Center for the Study of Immigrant Integration”. Southern California University Press.

Phalet, K., and Swyngedouw, M. (2003). “Measuring Immigrant Integration: the Case of Belgium”. *Studi Emigrazione, 40:773-803.*

Picarra, N., and Gil, A. R. (2012). “Research for the EUDO Citizenship Observatory. Country Report: Portugal”. European University Institute.

Ou, D., and Kondo, A. (2013). "In Search of a Better Life: The Occupational Attainment of Rural and Urban Migrants in China". *Chinese Sociological Review, 46.1, 25-59.*

Quheng, D., and Gustafsson, B. (2014). "The *Hukou* Converters—China's Lesser Known Rural to Urban Migrants". *Journal of Contemporary China, 23/88, 1-23.*

Reichel, D., and Perchinig, B. (2015). “Reflections on the Value of Citizenship – Explaining Naturalisation Practices”. *Austrian Journal of Political Science 44 (1): p. 32–45.*

Rubin, D.B. (1977). “Assignment to Treatment Group on the Basis of a Covariate”. *Journal of Educational Statistics, 2, 1-26.*

Sargsyan, V. (2017), Treatment-Related Naturalization Premiums in Two European Countries: Evaluation and Comparison. *International Journal of Economic Perspective 11(2), p192-196.*

Sawyer, C., and Wray, H. (2014). “Research for the EUDO Citizenship Observatory. Country Report: United Kingdom”. European University Institute.

Schröder, C., and Yitzhaki S. (2017). “Revisiting the Evidence for Cardinal Treatment of Ordinal Variables”. *European Economic Revie,w 92: p. 337-358.*

Song, Y. (2016). "*Hukou*-Based Labor Market Discrimination and Ownership Structure in Urban China: A Double-Selectivity Approach". *Urban Studies, 53(8)*, 1657-1673.

Song, Y. (2014). "What Should Economists Know About the Current Chinese *Hukou* System?" *China Economic Review, 29, 200-212.*

Stadlmair, J., 2017b, "Which Policies Matter? Explaining Naturalisation Rates Using Disaggregated Policy Data". *Austrian Journal of Political Science 46 (1): 59–74.*

Steinhardt, M. F. (2012). "Does Citizenship Matter? The Economic Impact of Naturalizations in Germany". *Labor Economics 19:813–823.*

Stern, J., and Valchars, G. (2013). "Research for the EUDO Citizenship Observatory. Country Report: Austria". European University Institute.

Van Oers, R., De Hart, B., and Groenendijk, K. (2013). "Research for the EUDO Citizenship Observatory. Country Report: The Netherlands". European University Institute.

Vink, M. P., Prokic-Breuer, T., and Dronkers, J. (2013). "Immigrant Naturalization in the Context of Institutional Diversity: Policy Matters, but to Whom?" *International Migration 51, no. 5.*

Vink, M., and Rainer, B. (2013). "Citizenship Configurations: Analysing the Multiple Purposes of Citizenship Regimes in Europe." *Comparative European Politics 11 (5): 621–648.*

Waters, C. W, and Jimenez, T. R (2005). "Assessing Immigrant Assimilation: New Empirical and Theoretical Challenges". *Annual Review of Sociology 31:105-125.*

Wong, L. and Wai-Po, H. (1998). "Reforming the Household Registration System: A Preliminary Glimpse of the Blue Chop Household Registration System in Shanghai and Shenzhen". *International Migration Review, 32/ 4, 974-994.*

Wu, X., and Treiman, D., J. (2004). "The Household Registration System and Social Stratification in China: 1955-1996". *Demography, 41, 363-384.*

Xing, C. (2014). "Migration, Self-Selection, and Income Distributions: Evidence from Rural and Urban China." *Economics of Transition*, *22/3, 539-576.*

Zhang, H. (2010). "The *Hukou* System's Constraints on Migrant Workers' Job Mobility in Chinese Cities". *China Economic Review, 21, 51-64.*

Zhang, L., and Tao, L. (2012). "Barriers to the Acquisition of Urban *Hukou* in Chinese Cities". *Environment and Planning, A 44(12) 2883 – 2900.*

Zhang, Z., and Treiman, D. J. (2013). "Social Origins, "*Hukou*" Conversion, and the Wellbeing of Urban Residents in Contemporary China". *Social Science Research, 42.1, 71-89.*

Zhang, Z, and Wu, X. (2012). "Registration Status, Occupational Segregation, and Rural Migrants in Urban China." Hong Kong University of Science and Technology, Hong Kong SAR.