UNIWERSYTET WARSZAWSKI

WYDZIAŁ NAUK EKONOMICZNYCH

INTERNATIONAL MIGRATION AND THE CHOICE OF SELF-EMPLOYMENT

Ph.D thesis submitted by Joanna Nestorowicz

Promoted by Professor Barbara Liberda

June 2013

Contents

Chapter 1

Introduction

International economic mobility will continue as long as differences between economies exist, be they in the domain of markets or in the domain of societies. This is in fact one of the oldest features of human civilization. In 2010 it was estimated that around 214 million people lived outside of their country of birth (United Nations Population Division, 2010). This state is far from being steady, though. Thousands of people per day make decisions concerning their place of residence and place of work. Opposite to what economic models and empirical verifications would suggest, the choices concerning location and form of economic activity can be made simultaneously. Labor market literature tends to treat migration choices as *status quo*, whereas migration literature is more concerned about migration *per se* than about the forms of migrants' economic activity. The resulting gap between the two approaches leads to oversimplified conceptual frameworks, which typically yield biased estimates and, consequently, inadequate conclusions. It is the objective of this thesis to partially fill in this gap.

The role of migrants in sending and receiving countries is in many cases vital. From the receiving country's perspective, immigrants satisfy the demand for labor, enhancing the productivity not only in the sectors where they work, but also in complementary industries. Furthermore, immigrants are usually net contributors to the host countries' tax systems. They can also be considered natural innovators. From the sending country perspective, emigrants often provide financial support to their family members left behind, thus facilitating investment in physical and human capital. On top of money, emigrants also remit cultural norms and knowledge which facilitate the implementation of innovative codes of conduct concerning both individuals or organizations. Indeed, all these benefits of migration are particularly strong, when migration meets entrepreneurship. The small and medium size enterprises often generate a considerable share of gross domestic product. While they rarely contribute to official R&D statistics, they constantly introduce novel ideas in terms of products and processes. Given the importance and benefits of migration and entrepreneurship for an economy, the combination of both can bring even greater advantages. Recognizing the potential value of these processes, the focus of this thesis concerns the nexus between migration and self-employment.

Migrant entrepreneurs are among those who are especially capable of taking advantage of the opportunities created by diverse market structures across countries. They can take account of capital (financial, social, cultural, physical) acquired at home and relocate it to a destination where it can generate higher returns. Such a form of arbitrage has been exploited since the Neolithic Revolution when populations which settled in various locations became able to generate surplus food production and get involved in trade. The new socio-economic order enabled distinguishing those who were immobile from those who became migrants. The link between migration and entrepreneurship, encompassing but not limited to trade activities, has existed ever since.

Conceptually, common pull factors suggest a natural link between migration and self-employment. Entrepreneurs are *par excellence* recognized as those who undertake risks. In fact, the very specificity of self-employment as a labor market strategy arises form time-, effort- and risk-demanding character. Migration is also a strategy associated with high costs and uncertainty of outcomes. It is also pursued with the aim of obtaining higher earnings. In this sense migrants are inherently entrepreneurial when compared to non-migrants, *ceteris paribus*. These arguments make immigrant self-employment conceptually a very coherent subject of study.

Notwithstanding, migration typically concerns citizens of poorer countries moving to richer countries. Under such conditions, migrant entrepreneurs are not likely to have an advantage: neither in terms of physical capital or technology, nor good knowledge of the local market conditions, regulations, consumer preferences, etc. In fact, one should expect that operating a business is harder for an immigrant than for a local citizen, *ceteris paribus*. Thus, immigrant entrepreneurs, if successful, typically perform arbitrage on the diversity of market structures and consumer preferences, exactly where it would be hard or nearly impossible under autarky and among solely autochthon communities. in light of these realities, this thesis will seek answers to the question of **what are the mechanisms through which migrants can take advantage of self-employment as a means to improve labor market outcomes**.

Theoretically, there are three primary mechanisms driving the choice of self-employment among migrants. First, migration and self-employment are in principle induced by differences in potential earnings. Second, some features of the host markets - e.g. discrimination - may prevent migrants from fully realizing their potential as employees, thus making self-employment an attractive alternative. Finally, the presence of co-ethnics may yield positive spillovers for immigrant business owners. Notably, the prevalence of discrimination can make immigrant self-employment an income-maximizing choice even if the migrants' (endowment adjusted) earnings remain below those of the native population. The positive spill overs and lack of entry barriers, however, could determine the relatively high profitability of self-employment in the hosting economy.

Each of the mechanisms outlined above generates incentives affecting international migrants in their choice of labor market activity. The incentives considered in this thesis are both direct and indirect. Based on existing migration theories, the difference in ex-

pected earnings between the receiving and sending countries will be treated as a direct incentive. The indirect incentives are related to two institutions which affect immigrants' business operation - discrimination and ethnic economies. Out of these two labor market features discrimination effectively "pushes" immigrants into self-employment as a second-best strategy. In contrast, ethnic economies induce favorable conditions for immigrant business in certain markets, ultimately "pulling them" into self-employment.

In the light of enormous cross-country differentiation in compensations, international mobility and mobility across labor market statuses may in fact enable workers to overcome the problem of human capital underutilization, understood as the under-compensation of individual productivity. On one hand, assets such as ethnic ties may increase the returns to self-employment among co-ethnic entrepreneurs, a feature existent only in the context of international mobility. On the other hand, strong links between the extent of wage discrimination and the intensity of self-employment seem to suggest that at least some of the self-employed originally intended to be wage-employed at the destination and it was only the change in labor market status that helped close the underutilization gap. Hence, this thesis formulates the following general hypothesis

General hypothesis *Self-employment, on top of international labor migration, facilitates reducing the inefficiency of the international division of labor.*

In search for proof of this proposition the thesis contributes to the existing literature by merging two strands of work. On one hand, it relies on the concepts and methods related to the economics of migration. On the other hand, the developments in the field of self-employment economics are taken into account. The migration literature either considers migrants as moving between two economies regardless of their labor market status, or, if it does deal with migrants' labor market choices, than only in the destination country context. The self-employment literature is concerned about the changes in labor market status, but of minorities in general, irrespective of whether one is a migrant or not. In the first case migration is accounted for, but not labor market status. In the second case labor market status is accounted for, but not the differences in market conditions between home and abroad. In the third case the specificity of one being a migrant is neglected. By bringing together the migration and the self-employment literature this thesis provides a refined understanding of immigrant self-employment, as a process which comprises both the choice of (labor) market and the choice of employment status on that market. On one hand, such knowledge will provide insights into whether facilitating migration among entrepreneurs can lead to more efficient allocations of capital. On the other hand, we will be able to determine whether it is worthwhile in terms of economic outcomes to support migrants in becoming entrepreneurs.

In order to derive conclusions on how gainful is the decision to become an immigrant entrepreneur, on one hand, earnings form self-employment abroad have to be compared to foregone earnings at home. In this case the extent to which immigrant self-employment

is an income maximizing choice depends on the pecuniary and non-pecuniary costs of mobility. On the other hand, the ability to determine the profitability of immigrant entrepreneurship also depends on the alternative costs of self-employment on the foreign labor market. Therefore, following the general hypothesis several operational hypotheses have been formulated in order to enable testing the specific mechanisms related to immigrant self-employment considered in this thesis. The motivation for formulating all of the hypotheses was a blind spots in existing literature. The question of whether immigrants earn more when they become self-employed abroad, then what they earn if they had not migrated will underlie the first empirical analysis. The operational hypothesis on this subject states that

The income gain hypothesis *Income from self-employment abroad is greater than expected income earned at home.*

Theoretical concepts of migration suggest that differences in earnings are an important stimulant for migration. Whether this holds for self-employed immigrants has not been empirically verified. The main challenge is methodological and lays in the ability to control for the selectivity of migration. The contribution of this chapter is, first, in taking advantage of the propensity score matching technique in order to overcome the selection problem. Second, the conducted analysis provides answers not only to the question of whether self-employed migrants earn more than what they would earn at home on average, but it differentiates between various labor market statuses in the country of origin. One of the reasons why the hypothesized income gain to immigrant self-employment could be especially advantageous, is that it enables overcoming labor market constraints in the destination country, such as employer discrimination. Assuming that discrimination is carried out on the basis of racial or ethnic group membership, such a context would be insignificant in the migrants' country of origin. In order to test for this scenario the following hypothesis has been formulated

The discrimination hypothesis *Labor market discrimination is related to increased intensity of immigrant self-employment.*

This above hypothesis is tested by applying a research strategy which quantifies the extent of discrimination by means of the Oaxaca-Blinder decomposition and then applies that measure in explaining the probability of immigrant self-employment. The main contribution of this study is acknowledging that the interrelation between discrimination and intensity of self-employment on local labor markets may vary in relation to various groups which comprise more or less "significant others" for immigrants. This conceptual novelty which considers a number of possible reference groups against which migrants may compare their earnings is especially relevant, as we are not interested in discrimination *per se*,

but rather in discrimination as a factor affecting the extent of self-employment activity among immigrants.

The migration context provides not only labor market limitations, such as discrimination, but also opportunities which are inexistent at home. An institution which creates such opportunities is that of an ethnic economy (enclave). The idea behind the functioning within ethnic enclaves is that one's co-ethnics may constitute a pool of consumers and labor beneficial for business operation. Furthermore, as this thesis will argue, the more co-ethnics take advantage of such a preferential setting and become entrepreneurs, the greater the competition, which is expected to be detrimental to profits. The hypothesis for this analysis is thus the following

The ethnic economy hypothesis. *Ethnic competition is related to lower, while ethnic complementarity is related to higher returns to local entrepreneurial activity.*

Considering both the opportunities and the threats related to the functioning of businesses within ethnic enclaves has not been done in existing research. This chapter contributes to the literature by introducing the Hirshman-Herfindahl Index of industrial concentration as a measure of competition computed for all entrepreneurs, and for immigrant entrepreneurs only. By comparing the estimates of the relationship between these two measures and the returns to immigrant self-employment we draw conclusions concerning the extent of exposure of immigrant business to general competition and to ethnic competition. In addition, we also control for the ethnic complementarity, by accounting for the size and the wealth of the co-ethnic community.

The underlying assertion of the latter two hypotheses is that the "foreignness" of immigrants may be both a handicap, as in the case of discrimination, and an asset, as in the case of ethnic economies.

Discrimination is typically hypothesized to push immigrants into entrepreneurship. Unjustifiably lower compensations from wage employment lower the alternative cost of self-employment, making it a profitable strategy at the margin. It should be noted, however, that often what is perceived as discrimination does not necessarily have to be discrimination. Different educational backgrounds and non-transferability of some skills may in fact imply an inability to discount some characteristics in the context of a foreign labor market. One such example may be high skills in the mother tongue coupled with relatively lower ability to communicate in the local language. For instance, highly skilled and talented actors may be unable to obtain (sufficient) recognition abroad due to a strong accent. In such circumstances, relatively lower compensation may in fact reflect lower endowments, not necessarily lower valuation of these endowments. Thus, although in principle discrimination could be an important push factor, it is also possible that even in "perfect" labor markets immigrant would prefer to be self-employed. Summarizing, immigrants' opportunity costs of self-employment may be relatively low or relatively high, what calls for an empirical analysis.

Notwithstanding, relatively high returns to self-employment due to positive spillovers from the co-ethnic population create the possibility for immigrants to entirely overcome discrimination in the hosting labor market. On one hand, immigrant entrepreneurs can be in possession of a comparative advantage resulting from home country specificity and employ it in a place where it is more scarce, thus potentially also generates greater returns. A good example are Indian immigrants employed in garmenting in Europe. Another type of comparative advantage may originate from a unique ability to satisfy the needs of the local (possibly co-ethnic) community, providing goods and services no other group can secure at comparable conditions. A good example here is the provision of religious services, dedicated foods, or dedicated food processing. With the growing size and welfare of the co-ethnic community, the range of such products and services expands, including familiar merchandise and tailored assistance. On the other hand, a larger number of co-ethnics also possibly implies better employment opportunities and stronger competition in such markets, what naturally lowers relative returns to business operation. Thus, both complementarity and competition between members of immigrant populations might affect the returns to immigrant self-employment. Furthermore, since immigrants tend to cluster geographically, it is the local density of co-ethnics that should drive the results. That being so, in case of sizable presence of immigrants in certain markets both relatively high and low returns to immigrant self-employment are theoretically possible. Determining which of these effects is quantitatively dominant requires empirical investigation.

The three aspects of immigrant self-employment mentioned above, i.e. the earnings disparities based on which immigrants can perform arbitrage, discrimination on the host labor market as a push factor, and ethnic complementarity as a pull factor into immigrant self-employment are the focal aspects of this thesis. However, a variety of local conditions in receiving countries make it empirically impossible to determine the quantitatively dominant pattern. As people from numerous countries go to various destination countries, their motivations and opportunity costs of mobility and self-employment will vary as well.

Figure 1.1a depicts the differences in immigrant and native self-employment rates across OECD countries. Though the OECD average shows quite similar proportions of immigrants and natives in self-employment, there are large regional discrepancies in the considered ratio. On one end of the spectrum, immigrants seem to be more entrepreneurial than the local population in Central and East European (CEE) countries. On the other end, in South European countries, it is the natives who are relatively more entrepreneurial. These differences may arise from factors differentiating the countries of origin as well as from factors differentiating the destinations.

The figure fails to depict the selectivity of home countries with respect to the considered host economies. While immigrants in the US or Australia come from all parts of the world, for most receiving countries it is possible to recognize a dominant sending state. Depending on the historical, social, political and economic differences between the home and the host economies, immigrants will experience varying modes of incorporation resulting in the observable diversity of labor market outcomes. These modes of incorporation can be

treated as origin-specific responses to the conditions immigrants are faced with at the destinations.

Behind the patterns represented in Figure 1.1a there are also dissimilarities in terms of the destination countries' market environments, e.g. the ease of doing business or the extent of market saturation in certain industries. Although these aspects might be considered universal, i.e. affecting both immigrants and natives alike, as emphasized a number of times heretofore, immigrants differ from the natives in their interactions with local economies. For example, in net emigration countries of CEE migrants will be differently perceived than in net immigration countries of Southern Europe. If they are considered as a threat (a common misinterpretation) it could be due to their abundance in the latter, or due to scarce information on their beneficial role in the economies of the former. The CEE markets will also have different institutional conditions for immigrant incorporation, both in terms of relatively underdeveloped legislation and organizational support for newcomers. These two features could affect the extent of labor market imperfections effectively targeting immigrants, such as labor market discrimination, and consequently also immigrants' self-employment rates.

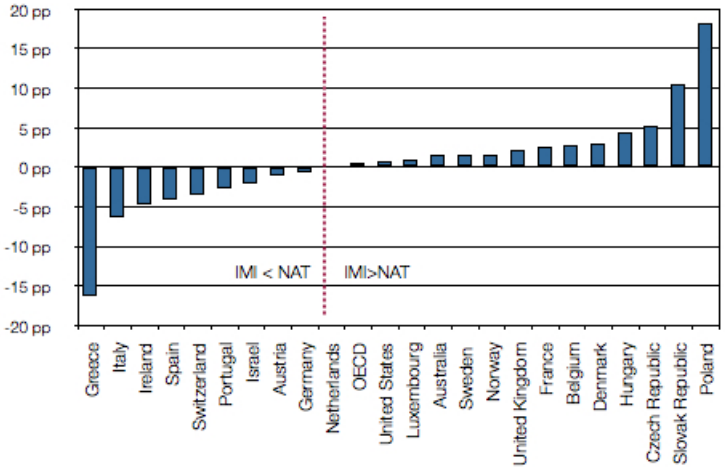
The various receiving countries also differ in the opportunities for immigrant entrepreneurship created by immigrant populations themselves. In the context of ethnic competition and complementarity relatively small immigrant populations may be a liability, as they limit the demand, or an asset, as they are unable to saturate the market. Figure 1.1b captures this aspect by plotting the differences in self-employment rates of immigrants and natives against the size of the immigrant population at a given destination.

The US stands out in Figure 1.1b as an especially interesting case. First of all, it hosts a very large number of migrants. This feature makes considerations concerning discrimination, but especially competition and complementarity effects very meaningful. Second, US immigrants and natives have very similar self-employment rates. Such conditions make research on immigrant self-employment more reliable in terms of possible generalizations of obtained results, as we are dealing with a setting which not extreme, but rather average.

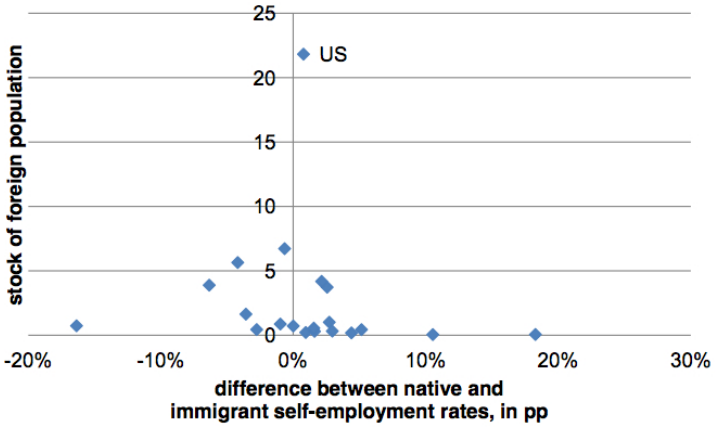
Even when considering only one destination country the diversity of the environments in the sending countries could easily masque the relationship between the alternative costs and the migrants' choice of self-employment. In order to compare choices of people who face comparable opportunities it would be preferable to consider a single sending country as well. In order to facilitate the measurement of the alternative costs in the sending country, it would also desirable to dispose of data concerning "stayers" of comparable quality to data at the destination. Finally, from an analytical perspective it would be favorable if the sending country under analysis was *relatively less developed* than the host economy. Otherwise, external validity of the findings would be limited.

One case that fulfills these requirements is Puerto Rico. In per capita terms, Puerto Rico is poorer than the US. In addition, unlike many other nations, Puerto Ricans do not face formal constraints concerning entry to the US, neither are they *de iure* limited by labor market regulations (e.g. they do not require work or residence permits in the US).

Figure 1.1: Relative immigrant self-employment prevalence across OECD countries.



(a) Difference between immigrant and native self-employment rates (in percentage points, circa 2008). Source: own elaboration based on OECD data: http://dx.doi.org/10.1787/888932440698.



(b) Difference between immigrant and native self-employment rates (in percentage points) vs stock of foreign population, circa 2008. Source: own elaboration based on OECD International Migration Database and http://dx.doi.org/10.1787/888932440698.

Furthermore, Puerto Rico uses the US dollar as its currency, what facilitates comparisons of earnings in the sending and receiving countries. Finally, Puerto Rico conducts a census survey according to the same standards as the US, what limits the scope of potential mismatch between the measures used for the sending and for the receiving state.

Given the advantages, this thesis will analyze primarily the situation of the Puerto Rican "stayers" and emigrants to the US. Such individual case studies, however, often suffer from interpretational limitations in terms of external validity. Thus, for each of the analyses, we attempt to provide a generalization against entrepreneurial choices made by people who face comparable opportunities.

Summarizing, this thesis contributes to the literature in three major ways. First, it combines the labor market literature and the migration literature from a new perspective.

All three notions - the expected income gain, the discrimination, and the influences induced by one's co-ethnics - are present in existing works, but not in the contexts considered in this thesis. Income gains, though analyzed for wage-employed migrants, have not been researched for self-employed migrants. Discrimination, in spite of having been examined among immigrants in general, has been rarely looked into as a feature affecting immigrant self-employment rates. Finally, market competition and complementarity are often studied in relation to the mainstream economy, but are still an under-researched topic, especially when considered jointly, in the context of ethnic economies.

Second, we provide a number of policy relevant findings. First, this thesis shows that there is a significant income gain to being a self-employed immigrant - both in reference to wage-employment abroad and self-employment in the home country. Second, it proves that, indeed, discrimination (even in relative terms) can be an important push factor for self-employment. Finally, although competition generally lowers returns to self-employment, in the case of immigrants, it must not, while the co-ethnics' larger purchasing power generates additional stimulus for profitability.

When taken together, these findings shed new light on the international earnings disparities. In fact, it seems that *via* self-employment individuals are able to considerably narrow the size of the international compensation gap. On one hand, migration is typically treated as a means for reducing the under-utilization of the human capital, especially for migrants from less developed countries. On the other hand, labor market discrimination against immigrants justifies policy interventions aimed at facilitating immigrant inclusion in the host countries. This thesis demonstrates that immigrant self-employment not only fulfills the expectations of migration, i.e. the effective reduction of the under-utilization of migrants' human capital, but moreover, that it leads to more effective allocations of human capital, enhancing immigrants' labor market outcomes at the destination and generating multilateral benefits for migrants and the receiving societies alike. While typically inclusion policies stress the role of adaptation, skills, and open access to wage employment, this thesis suggests that policies facilitating self-employment among migrants may in fact be more effective in limiting the under-utilization problem.

Last but not least, we contribute to the methodology of migration and labor studies by adapting the available empirical techniques to test the formulated hypotheses. While propensity score matching, parametric decomposition as well as returns and choice models have been used extensively in research in general, our applications are novel in each of the relevant strands of literature. The application of propensity score matching does not aim to provide *causal* interpretation, but to address the selection bias problem. Since it is applied to relatively large samples in a relatively homogeneous context, the validity of such a method choice seems satisfactory. Second, the estimates of discrimination are used as explanatory variables in our framework, rather than the end point of the analyses. A number of robustness checks as well as external validity exercises suggest that the results are fairly general and statistically strong. Finally, we are the first to include

both complementarity and competition measures, when analyzing the returns to self-employment.

This thesis comprises an extensive literature review, three independent empirical analyses and a concluding section. Indeed, is immigrant self-employment a beneficial choice? This thesis will analyze predominantly Puerto Rican immigration to the US. In Chapter 3 we briefly describe the characteristics of Puerto Rico and the US, which justify such choice of empirical material. The following chapter presents an overview of migration, self-employment and immigrant self-employment theories and theoretical concepts. In accordance with the structure of following, empirical chapters, the literature review is divided into three parts. The first part discusses the determinants of migration. The second part provides an overview of the determinants of self-employment. The third part considers the two selected contexts of immigrant self-employment which will be analyzed in this thesis, i.e. discrimination and ethnic economies. While there are numerous reviews of both migration and self-employment literature, our objective was to show the main existing knowledge gaps. In addition, it organizes general theoretical foundations for subsequent analyses. This thesis is predominantly empirical, putting a lot of stress on the methodological context of migration-cum-entrepreneurship studies. However, methods employed in each of the three analyses differ substantially. To facilitate the flow of analyses, the review of the advantages and disadvantages of various empirical strategies is discussed in each chapter separately, along with the main insights form earlier empirical works.

Three empirical chapters follow the literature review. Each of them is dedicated to testing specific operational hypotheses. The last section of each empirical chapter provides an extension of the relevant study. The extensions allow to consider the analyzed determinant of immigrant self-employment, be it income differences, discrimination or ethnic competition and complementarity, in a wider context, i.e. one not limited to Puerto Rican immigration to the US.

The first chapter tackles the most fundamental issue of income gains to immigrant entrepreneurship. It provides methodologically robust estimates of the counterfactual earnings of emigrants, which would have been gained had they not migrated. We employ propensity score matching to address the problem of migration (and entrepreneurship) selectivity. We compare income from self-employment in the receiving economy to income form wage employment in the receiving economy and to income from self-employment and wage-employment in the sending economy. The core of the analysis concerns self-employed Puerto Rican "stayers" and immigrants to the US. We find that in fact returns from self-employment in the US are superior to alternatives. The test of external validity is provided by an analysis of self-employed "stayers" and migrants in the context of inter-state mobility in the US. Only relocations between states as distant as Puerto Rico is from the US are considered in order to enable comparability of the results.

The second empirical chapter deals with labor market discrimination as a reason for undertaking self-employment by migrants. Based on the Blinder-Oaxaca decomposition

of wage differences it estimates the probability of becoming self-employed under the conditions of local labor market discrimination. In fact, the analysis comprises two parts. First, the Oaxaca-Blinder decomposition is applied to wage equations of Puerto Rican immigrants in the US and several comparison groups of natives. This stage of the research allows us to determine the extent of labor market discrimination. Second, the estimates of discrimination obtained in the first stage are incorporated into probability models of immigrant self-employment, allowing to recognize the interdependencies between discrimination and the intensity of immigrant self-employment. The findings prove that the extent discrimination is positively related to self-employment incidence among Puerto Rican immigrants. The external validity of the study is tested on several aggregate ethnic groups of immigrants residing in the US.

The third empirical chapter analyzes the hypothesis of ethnic enclaves as potential stimulants for immigrant self-employment. The innovative approach introduced in this chapter considers not only the potentially positive effect of ethnic congestion, though, but also the possibly detrimental effect of ethnic competition. The former is expected to increase profits by enlarging the customer and labor pool. The latter is expected to decrease profits by reducing the market power of a representative firm. Ethnic complementarity is primarily operationalized as the share of co-ethnics in a local market. The co-ethnic group's relative wealth is also controlled for. Ethnic competition is operationalized as the value of the Hirschman-Herfindahl Index. We compute it not only for co-ethnics, however, but also for the whole population of entrepreneurs. The results show that although ethnic competition is detrimental to profits, increased external competition provides a favorable environment for ethnic businesses. The effect of ethnic complementarity as such was found unrelated to business profits, but increased purchasing power of co-ethnic communities was positively related to the returns to ethnic entrepreneurship. The external validity of the study is provided by an analysis of the competition and complementarity mechanisms among the same aggregate ethnic groups as considered in the extension to the previous chapter.

The last chapter develops policy implications, summarizes the findings, and provides suggestions for future research.

Chapter 2

Literature review

This chapter aims at demonstrating gaps which exist in the literature on migration and self-employment. It shows that immigrant self-employment, as an outcome of a migration decision and a labor market status decision, should be an income maximizing choice. It further establishes that, on one hand, self-employment is a form of labor market activity which enables migrants overcome labor market imperfections such as discrimination. On the other hand, it also determines that there exist contexts such as the ethnic economy, which enable self-employed immigrants taking advantage of their joint migration and labor market status.

The presented discussion reveals the drawbacks of existing research and substantiates the theoretical and methodological approaches applied in this study. The chapter is divided into three parts, each dedicated to a specific notion. The first two parts consider analyses related to the determinants of labor migration and the determinants of self-employment as forms of labor market activity. The final section takes the choice of migrating and becoming self-employed as given and reviews the theoretical concepts related to two contexts of immigrant self-employment - labor market discrimination and ethnic economies.

In this chapter only theoretical issues are tackled. Reviews of methodologies applied in testing these theoretical predictions are discussed in each of the empirical chapters of the thesis.

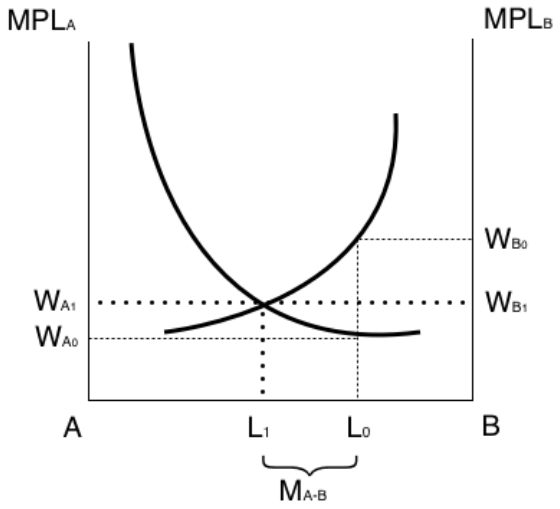
2.1 Determinants of labor migration

There is no holistic, unified theory of labor migration[1]. The multitude of individual preferences and objectives regarding (international) migration, together with the numerous contexts in which they are translated into actual strategies lead to the evolution of various theoretical concepts explaining why mobility occurs. Each one of them tackles only a selection of issues and overlooks the influences of the others. This section elaborates on a catalogue of most widely recognized factors which affect migration choices.

[1]Due to its labor market orientation this thesis focuses solely on labor migration and excludes forced mobility or that relying on other motivations, such e.g. as educational.

All seminal concepts in the economics of migration consider expected income differences as one of the determinants of migration. Hicks (1932) was the first to make a very straightforward reflection that "differences in net economic advantages, chiefly differences in wages are main causes for migration" (Hicks, 1932, p.76). This statement has become the motto of all neoclassical migration theorists. The idea of differences in net returns to economic activity on two markets as the main driver of mobility has been most often exploited in the framework of trade theory. Bhagwati and Srinivasan (1998) have proposed labor migration to be considered in the same way as the mobility of capital. Flows between markets were expected to take place until wages equalize. A graphical representation of their approach has been replicated in Figure 2.1.

Figure 2.1: International trade model of labor mobility.



Source: Bhagwati and Srinivasan (1998).

In T_0 employment in countries A and B is AL_0 and BL_0, respectively. Country B, where labor is relatively more scarce, offers relatively higher wages ($w_{B0} > w_{A0}$). Neoclassical theory claims that this creates a sufficient incentive for people to want to move from the lower paying market (A), to the higher paying market (B). The model predicts that as long as wages are strictly higher in B than they are in A, people will move. As they move labor becomes more and more abundant in country B and more and more scarce in country A, decreasing marginal productivity of labor in the former (MPL_B) and increasing it in the latter (MPL_A). An equilibrium with equal wages is reached when M_{A-B} workers migrate from A to B, at which point $w_A = w_B$ leaving no more incentive for mobility.

Despite its simple and convenient mechanics, the neoclassical model of migration has been considered an oversimplification for at least three reasons. First, due to institutional, technological and structural factors equalization of wages does not necessarily occur. The poorest countries in the world do not become completely depopulated and migration flows between some very diverse states are inexistent or at least much less intensive than expected given the extent of wage disparities. The second very much related oversimplification lies in that labor is not homogenous. People actually may have different propensities

to migrate. This associates the choice of being mobile and of choosing a given labor market status to individual characteristics, rather than just macroeconomic features of markets. The third reason why there exists a mismatch between neoclassical migration theory and reality is the existence of alternative stimulants which significantly affect the decision to migrate, even in case of wage equality between the sending and receiving economies.

In an attempt to refine the neoclassical approach to migration, Sjaastad (1962) proposed that migration be treated not as a process subject to market equilibrating mechanisms on the macroeconomic level, but as an individual resource allocation decision. As such, migration in Sjaastad's (1962) view involves both monetary costs, including the opportunity cost of earnings forgone during the migration process, and non-monetary costs, i.e. "psychic costs" (Sjaastad, 1962, p.85). This understanding of the individual migration decision-making process implies that a potential migrant aims at maximizing the present value of lifetime net gains to mobility. Such a view of migration can be formally expressed as an individual's decision resulting from maximizing the following equation (Bodvarsson and van der Berg, 2009, p. 33)

$$\pi = \int_{t=0}^{T} [W_t^M - W_t^H - CL_t^M + CL_t^H] e^{-rt} dt - C(D, X).$$

The gains are modeled as a function of: the, potentially foregone, earnings at home (W_t^H); the "costs" of living at home, which can be treated as non-pecuniary benefits of remaining in one's place of origin (CL_t^H); the, potentially gained, financial benefits from migration (W_t^M); the costs related to finding oneself in a strange context, i.e. the costs of living abroad (CL_t^M); other fixed costs related to distance (D); and personal characteristics, preferences etc. (X). The main drawback of Sjaastad's formulation of a migration decision was that it assumed that an immigrant would be able to be certain about finding a job at the destination. It could also be criticized for not being able to capture multiple migration events in one's lifetime and for not including the migrant's household's role in the decision process. Given its formula, which considers migration as a one time occurrence, Sjaastad's understanding also fails to account for the benefits (to the household, thus partially also to the migrant) and the costs (to the migrant) of sending remittances.

Harris and Todaro (1970) reformulated Sjaastad's proposal in a model which included a probability term in the income maximization problem. Based on an analysis of rural-to-urban migration they have observed that people migrate despite high unemployment in the modern sector in the cities. An explanation of this phenomenon was found in the expected income hypothesis, which stated that people will migrate as long as their expected wage in the urban sector, net of their migration costs, is higher than what they would earn in the countryside.

Sjaastad's (1962) and Harris and Todaro's (1970) models of migration recognized more than just the pecuniary gains and costs to migration. In this sense they ask the question whether migration is a consumption decision or a supply decision (Bodvarsson and van der Berg, 2009). Treating migration as more than just an optimal allocation of labor necessitates considering other gains to migration than higher expected earnings. Better living

conditions, wide-ranging civic rights or even a more preferable climate have been considered possible gains of that sort (Sjaastad, 1962; Lee, 1966). Costs complimentary to such gains have also been recognized. Among them distance (Ravenstein, 1885; Sjaastad, 1962; Lee, 1966), political barriers (Lee, 1966; Harris and Todaro, 1970), or costs related to family separation (Lee, 1966) have been identified. In order to be able to explain why individuals do not respond to these factors in the same way, migration theorists have proposed that they are moderated by individual characteristics (Ravenstein, 1885; Lee, 1966; Sjaastad, 1962).

Selectivity of migration

The postulate that individual characteristics are a determinant of migration inevitably lead to formulating the presumption that migration is a selective process. Lee (1966) observed that age, gender, level of educational attainment, degree and strength of social ties, one's social class or cultural awareness may hypothetically differentiate the strength of responding to factors pushing and pulling migrants to and from certain locations. In extreme cases, as Lee suggested, certain personal features could either be catalysts for the migration process, or they could completely prevent one from engaging in mobility.

As labor migration is a labor market choice it goes without saying that it concerns people in productive age. Within that group those mid-aged (30-40 year olds) are in practice most likely to become international migrants (United Nations, 2011). Graduates may prefer to initially search for employment opportunities on the domestic labor market. Moreover, the decision to stay in the country of origin may also be dictated by the need to set up families. Older cohorts may choose to be immobile due to their health limitations and the harsh work conditions usually awaiting migrants abroad. If people do not have previous migration experience it may be also difficult or them to adapt to a new environment at a certain age. Gender is an aspect of migration strongly related to the demand side of certain migration streams. While flows from some countries tend to highly feminized, others are masculinized, depending on the type of jobs immigrants take up at the destination (DeJong, 2000; Pessar and Mahler, 2003; Massey et al., 2006). Familial arrangements may one hand be push factors as people search for the means to provide for their kin. On the other hand, they pull one to stay at home as the costs of family separation may be considered very high (Mincer, 1977; Dumon, 1989; Kofman, 2004). Education and one's level of human capital may also hypothetically work in both directions. On one hand, better endowment in human capital increases the probability of finding a job abroad, which should increase the propensity to migrate. On the other hand, it also increases one's probability to find a job at home, increasing the opportunity cost of working abroad.

Borjas (1988) has relayed on Sjaastad's neoclassical view of migration and formalized it in a way that took into account the heterogeneity of potential migrants and their varying responsiveness to the determinants of migration. In his model Borjas considered workers with varying distributions of skills between the sending and receiving societies. He also

assumed incomplete transferability of skills across markets. In Borjas's framework a person would migrate if he or she experienced a positive value of the index function, I, which takes the following form (Borjas, 1988, p. 4)

$$I = ln(\tfrac{w_1}{w_0+C}) \approx (\mu_1 - \mu_0 - \pi) + (\epsilon_1 - \epsilon_0)$$

where w are wages, C is the cost of migration, π is the so-called 'time-equivalent' cost of migration ($\pi = C/w_1$) and, following Roy (1951), μ are the observable wage determinants, ϵ are the unobservable wage determinants. Indices 1 and 0 represent the destination and source country, respectively. The probability of migration, i.e. the probability of $I > 0$ was positively related to mean income at the destination, negatively related to the mean income in the source country, and negatively related to the cost of migration. These predictions of the model are in line with what was proposed by Sjaastad (1962).

What Borjas's model was also able to represent, contrary to Sjaastad's approach, was that the type of migrants for whom it is feasible to emigrate to a given destination country depends on the ratio of the variances of the income distributions of both economies. If it was only the mean earnings that mattered, as Borjas argued, we would observe all the people from the "poorer" country migrating to the "richer" country, what clearly is not the case. Borjas illustrated his argument with a simple example of when σ^2, the variance of the unobservable determinant of wages, is higher is the country of origin than at the destination. Low-earning workers (implicitly also low-skilled) would then be "insured" against relatively poor labor market outcomes. For the high-earners (high-skilled workers) the situation would be disadvantageous due to the compressed distribution of earnings. The opposite would be the case if $\sigma_1 > \sigma_0$. In the former case the destination would create incentives for rather low-skilled workers to immigrate, in the latter case rather high-skilled workers would be encouraged to come. In his work Borjas neatly merged earlier hypothesis in the field of labor migration with established labor market theory. Despite the refinements to former models, Borjas's representation of migration processes remained in the neoclassical framework, assuming that migrants are rational agents with complete information, able to weigh the costs and benefits of mobility accordingly.

The New Economics of Labor Migration (NELM) formulated by Stark and Bloom (1985) was the most influential concept proposing to amend the dominant neoclassical paradigms. Stark and Bloom (1985) suggested that the rational migrant is not the sole agent deciding whether or not to migrate and that the migrants' social environment plays a critical role in the process. Based on these premises the authors were able to demonstrate that the differences in wages need not be a necessary condition for migration to occur.

In light of Stark and Bloom's (1985) theory migrants were embedded in a social environment composed of their family or household members and a constructed group of reference, be it their neighbors, peers, co-workers or others. Stark and Bloom thought of the familial group, rather than the migrant alone, as the main agent in the migration decision-making process. Building upon this advancement in thinking about migration Stark and Lucas (1985) have argued that the relationship between a migrant and his or

her family is a contractual one, what is visible e.g. in remittance patterns and treating migration as a household, rather than an individual, strategy. Another aspect, which enforced viewing migration as a collective, rather than individual decision, was that it could be treated as a risk diversification strategy for the household (Stark and Bloom, 1985). If migration of one or more family members took place to a country where the earnings were negatively- or un-correlated with the earnings in the country of origin (where other family members worked), then the cost of migration (including family separation) could have been considered as the cost of family insurance against negative shocks affecting employment in one of the markets. Relating notions such as risk diversification or within-family contractual agreements to migration was unthought of in the main streams of previous literature. Such thinking conceptually allowed for the occurrence of migration despite no significant differences in wages between the origin and the destination. As far as the constructed reference group is concerned Stark and Bloom (1985) suggested that people might be willing to migrate in order to change their relative social position or to change their group of reference. From the latter it follows that after migrating people would rather compare themselves to other migrants than to the members of the receiving society. This further allows to explain why migrants are willing to fill in niches unwanted by the natives and undertake dirty, dull and demeaning (3-D) jobs.

Summary

An overview of the most influential contributions to labor migration theory reveals two major dimensions of their scope. One considers the exogenous mechanisms which are hypothesized to frame people's choices. Here we find the theories which exploit equilibrating market mechanisms on one hand, and the numerous structural push and pull factors on the other hand. The second dimension of how migration processes are understood is related to the migrants' endogenous choice making mechanism, i.e. the form and components of his or her objective function. Whether the migrant maximizes solely income or rather a bundle of components of utility, including market and non-market consumption goods, remains a significant difference across various propositions.

Given the number of migration theories and the foundations for migration research they propose, the following features are recognized as most critical for the presented research:

- the notion of migrant selectivity,
- expected income differences as the main determinant of migration as a labor market strategy,
- the social context in which the decision to migrate is made with both its familial and other components.

Out of the catalogue of determinants of labor migration two will be purposefully set aside in the following research. The first is the notion of possible non-pecuniary costs and benefits of migration and the second one is the idea of risk-diversification as an alternative

migration strategy. Primarily both of these issues are very difficult to capture in the data. Moreover, setting aside benefits from migration which are of other than financial character is due to the fact that they are subject to individual preferences as to what should be considered a cost and what should be considered a benefit. Some people might prefer to live in a warmer climate, while others in a cooler climate; some like to have clear and strict rules about how to behave, while others would rather make their own decisions within a wide scope of acceptable practices. In this sense it is impossible to determine what should be the effect of certain circumstances on one's cost-benefit analysis of migration. Nevertheless, both pecuniary and non-pecuniary costs and benefits of migration will be considered in the interpretation of obtained results.

2.2 Determinants of self-employment

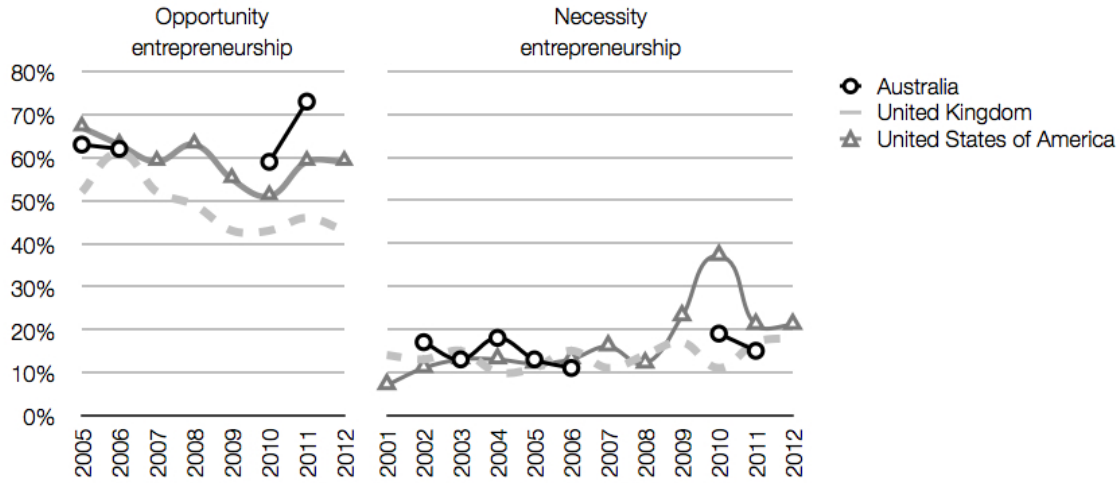
Similarly to the determinants of migration, the determinants of self-employment may also be recognized as individual and contextual, i.e. related to the socio-economic circumstances under which one undertakes entrepreneurial activity[2]. The former lead to the selectivity of self-employment. The latter may be considered in two major aspects - circumstances that create an opportunity for establishing entrepreneurial ventures and circumstances that push people into entrepreneurship as an option of last resort.

Among those who stress the need to differentiate between entrepreneurship out of necessity as opposed to entrepreneurship resulting from voluntary realization of a business plan are Earle and Sakova (2000), Llisterri et al. (2006), or Naude (2008). The rate of the two types of entrepreneurial choices across countries has been estimated within the Global Entrepreneurship Monitor project (GEM, www.gemconsortium.org). GEM defines the relative prevalence of "Improvement-Driven Opportunity Entrepreneurial Activity" as the percentage of those involved in early stage entrepreneurship who "(i) declare that they are driven by opportunity as opposed to finding no other option for work; and (ii) who indicate the main driver for being involved in this opportunity is being independent or increasing their income, rather than just maintaining their income" (Global Entrepreneurship Monitor, 2013). Relative prevalence of "Necessity-Driven Entrepreneurial Activity" is defined as the "percentage of those involved in entrepreneurial activity who are involved in entrepreneurship because they had no other option for work" (Global Entrepreneurship Monitor, 2013). The rates of opportunity and necessity driven entrepreneurship in major immigrant receiving countries are presented in Figure 2.2.

The US exhibited an upsurge of necessity driven entrepreneurship accompanied by slight decrease in the relative number opportunity driven startups in the years of the global financial crises. This pattern is not as clear in the case of the UK, where opportunity driven entrepreneurship falls since 2006 and no significant increase in necessity entrepreneurship is visible. Data for Australia is incomplete for that period. Unfortunately

[2]The review of existing literature presented in this subsection has been partially discussed in Tyrowicz and Nestorowicz (2011)

Figure 2.2: Scale of opportunity- and necessity-driven entrepreneurship.



Source: own elaboration based on the Global Entrepreneurship Monitor data, www.gemconsortium.org/Data.

GEM data do not allow to differentiate between native and immigrant entrepreneurs. We could speculate that, because the two groups face the same macroeconomic conditions, be they preferable for business or not, the motivations behind their labor market choices are similar. This might not be the case for two reasons. First, as it will be demonstrated in further sections of this chapter, immigrants and natives do not necessarily operate on the same markets. Second, immigrants are a rather disadvantaged group on the labor markets of receiving states. As these two features potentially have opposing effects, the former creating potential for more opportunity entrepreneurship and the latter for necessity entrepreneurship, it is impossible to determine how do immigrants fare in comparison to the natives in these aspects.

Whether opportunity- or necessity-driven entrepreneurship is on the rise in aggregate terms depends partly on the business cycle as such, but the local economic conditions affect people's entrepreneurial choices as well. Ferrante and Sabatini (2007) found a negative effect of regional *per capita* income on the choice of self-employment. Availability of less risky careers in wealthier regions bears a lower opportunity cost that becoming an entrepreneur. Other regional effects which may influence the probability of becoming self-employed are local labour market related factors such as the unemployment rate (Van Praag and Van Ophem, 1995; Blanchflower and Oswald, 1998; Audretsch et al., 2001; Baptista and Thurik, 2004), the local industrial organization structures (Malecki, 1993; Storper, 1995), or regional institutions and policies (Amin, 1999; McQuaid, 2002; Lee et al., 2004).

Irrespective of the global and local economic conditions and the resulting rates of opportunity and necessity-driven entrepreneurship, *ceteris paribus*, some people do engage in self-employment while others do not. This leads to exploring the selectivity of self-employment, i.e. the individual characteristics which serve as explanatory variables of entrepreneurial activity on the labour market. Among them gender is one of the most

widely discussed. Cowling and Mitchell (1997) claim that self-employment rates among women may be an effect of discrimination by the employers. An argument which will also be used in the context of immigrant self-employment. In this sense women are treated, similarly to foreigners, as a labor market minority. Other reasons behind the preference of women for self-employment are presented by Hughes (2003) who puts forward such issues as proactive responding to lack of other opportunities, more flexible working hours, the ability to work at home and remain a care-taker for children or older family members. Despite the preference for self-employment, fewer women than men might in fact be self-employed. Dolton and Makepeace (1987) attribute this observation to treating activity on the labour market as complimentary for house work, in which case involvement in the latter decreases activity in the former.

When the character of labor market activity is considered at the household level, marital status has also ben found to play a critical role in the decision to become self-employed. Taylor (1996), Fairlie (2005), Hyytinen and Ilmakunnas (2007) or Tamvada (2007) mention positive correlations between being married and the propensity of being self-employed. In their view marriage is said to reduce entrepreneurial risk through income diversification, given the spouse is economically active. Marriage may also possibly be a source of start up capital due to the ability of the couple to pool their resources. Le (1999) recalls a further number of reasons why there may be a positive correlation between being married and self-employed, especially when the spouse is also engaged in the business. In this case, apart from some of the possible reasons mentioned previously, the spouse increases the levels of human and social capital and provides reliable labour, the latter aspect being also analyzed by Borjas (1986).

On top of gender-related issues, human capital was also found relevant for explaining selection into self-employment. Cross sectional studies prove it is one of the most powerful variables used to explain the selection processes. Its role remains ambiguous, however (Dickson et al., 2008). Van der Sluis et al. (2004) claims that the differences in results obtained from various studies arise from diverging returns to self-employment from different levels of education. Analysis based on US and European data draw on a theory of how the schooling system creates incentives for self-employment. Within this stream of literature the role-modeling effect of parents' occupational choices has also been considered. Blanchflower and Oswald (2007) concluded from their quantitative analyses that the probability of becoming an entrepreneur increases if one's father has been self-employed. There is also evidence which demonstrates that supporting one's family members' entrepreneurial activity may partially compensate for formal education when one becomes self-employed (Silva, 2007; Hartog et al., 2008; Caliendo and Kritikos, 2008). In the case of immigrants, as it will be discussed further, cultural capital, which can be considered a form of human capital, also plays a critical role in the selection process.

Apart from gender, marital status and human capital also age has been extensively analyzed as a determinant of self-employment. Numerous studies have found positive correlations between age and the probability of becoming self employed in developed economies

(Brock and Evans, 1986; Rees and Shah, 1986; Borjas and Bronars, 1989). The selection process with respect to age does not follow a linear path, though. On one hand, representatives of the youngest cohorts would not get involved in setting up their own business due to lack of start-up capital and know-how. Neither the oldest cohorts in productive age would do so, because of risk aversion related to potential loss of life-time savings just before retirement. This type of thinking would lead to concluding that the greatest odds of becoming self-employed are among the mid-aged population. Having gathered enough financial and human capital and still having a long horizon of possible activity on the labour market in case of business failure, individuals in their 30s, 40s and 50s are said to be those who most probably become self-employed. On the other hand, bearing in mind the notion of necessity-driven entrepreneurship, we could theoretically assume that the least productive cohorts - recent graduates with no labor market experience and the eldest workers would be most disadvantaged, implying their increased propensity to undertake self-employment as a survival strategy (Llisterri et al., 2006).

Two other characteristics which hypothetically affect one's propensity to become self-employed are risk aversion and entrepreneurial ability. Risk aversion as a feature characterizing the self-employed has been recognized by Cantillon (1755). In his essay Cantillon defined the entrepreneur as one who performs arbitrage and bears the risk of uncertainty. Relatively low risk-aversion is considered immanent to entrepreneurship till now (Levine and Rubinstein, Levine and Rubinstein). Entrepreneurial ability has been conceptualized as the capability to coordinate factors of production (Say, 1803), introduce innovations (Schumpeter, 1947), manage human resources (Leibenstein, 1968), or possess specific psychological traits (Parker, 2006; Caliendo and Kritikos, 2012).

As difficult to measure as they are, entrepreneurial ability and risk aversion have been considered in theoretical models of entrepreneurial choice. Lucas (1978) considered a closed economy, with a fixed homogenous capital stock, and with workers who identical as far as their productivity in wage-employment was concerned, but differed in their productivity (managerial ability) in self-employment. It is assumed that higher ability translates into higher output regardless of the amounts of capital and labor employed in the production process. Individuals with lower ability than that of the "marginal entrepreneur", i.e. one for whom the net profits from self-employment are equal to the going wage, remain wage-employed. All those with better ability choose self-employment. Demand functions derived from Lucas's model indicate that the firms' demand for capital and labor are increasing in the managerial skills of individuals running them, implying that more able managers run larger firms.

Lucas's model explains occupational choice between wage-employment and self-employment, but the choice is assumed to be made once in a lifetime and risk related to undertaking an entrepreneurial venture is ignored. Dixit and Rob (1994) suggest a two-sector model which incorporates uncertainty into the production of output in one of the sectors, thus enabling conceptually treating it as if it was a (risky) "self-employment sector". Output in the other sector is certain, as if it was the case of wage-employment.

Individuals are able to switch between the sectors, i.e. forms of labor market activity. Switching is costly. The agents are risk-averse and impatient. Thus the agents maximize their discounted expected lifetime utility net of the costs of switching

$$E\left\{ \int_0^\infty U[y(t)]e^{-\delta t}dt - \sum_i ce^{-\delta t_i} \right\}$$

where U is the utility, y is income, c are the costs of switching, while t_i indicates the periods when switching is done. The implication of Dixit and Rob's (1994) model is that there is a certain (high) threshold level of output in self-employment which stimulates agents to switch from wage-employment to self-employment. There is also a (low) level output which lowers expected utility to the extent that individuals switch from self-employment to wage-employment (where output is certain). If output in self-employment falls in between these two thresholds individuals remain in their initial form of labor market activity, as the costs of switching would make them worse off. The probability of switching increases as risk aversion falls, as impatience rises, and as switching costs become lower.

Summary

The main findings of the self-employment literature provide quite clear pointers as to the determinants of entrepreneurial labor market choice. These determinants can be classified into individual and contextual. Among demographic variables we find age and gender. Other features than gender, such as race or ethnicity, which would classify one as a minority are also considered. One's marital status and level of educational attainment have also been found to be significant correlates of the propensity to become self-employed. Neither of these characteristics enables making unambiguous predictions as to the directionality of their relation to entrepreneurial behavior, though. The reason lies in the contextual factors which define the role of certain characteristics in a given economic environment. It makes a difference whether one is forced to become self-employed due to unfavorable conditions on the local labor market, or is inspired by the existing opportunities to become an entrepreneur. The following section will discuss in greater detail two contexts, one potentially "pushing" and the other "pulling" into self-employment. The specific case of immigrants as labor market actors will be considered.

2.3 Contexts of immigrant self-employment

Literature concerning immigrant self-employment has combined features established in the field of migration with those established in studies of entrepreneurship[3]. From the former it takes the specificity of migrants in cultural terms or the role of networks in the migration process. From the latter it mainly adopts the ideas of "self-employment out of necessity" and the basic premise of entrepreneurs being market actors who perform arbitrage.

[3]This section is based on a wider overview of immigrant self-employment theory published as: Nestorowicz, Joanna (2012). Immigrant Self-employment: Definitions, Concepts and Methods, Central and Eastern European Migration Review 1(1).

In the context of performing successful arbitrage immigrants are inherently entrepreneurial. The fact that they strategically relocate their human capital to a country where it generates higher returns makes them exemplary entrepreneurs. Self-employed immigrants have an advantage over their wage-employed counterparts in that they are not as dependent on the labor demand at the destination country and can take advantage of their social and cultural capital in order to maximize gains to migration. These specific features are most visible in two contexts of immigrant self-employment. The first is labor market discrimination. Becoming self-employed detaches one from the unfavorable labor market circumstances and possibly allows to obtain relatively fair returns to one's productivity. The second context is the presence of other immigrants on a local market. Whether there are many co-ethnics at a given destination may be a necessary condition for the success of an immigrant business venture. It enlarges the pool of consumers and the supply of employable labor. In the light of the approaches to the choice of self-employment presented in the previous chapter we may say that discrimination is a factor which pushes immigrants into necessity-driven self-employment. High concentrations of one's co-ethnic may be an opportunity which allows to make the choice to become self-employed profitable. In the following two sections theories related to the above-mentioned contexts and the mechanisms through which they affect immigrant self-employment will be discussed.

2.3.1 Labor market disadvantages

The economic theory of discrimination focuses on unequal labor market opportunities of "minorities", e.g. women, disabled persons, or people representing certain racial groups. The stream of work stems from the economic writings of Becker (1971) and theoretical models of Arrow (1998), who provide insights into how race may provide signaling effects on the labor market, serving as a screening device for employers, who extrapolate conceptions concerning e.g. productivity to the whole ethnic groups. This exclusion from certain (or all) sectors of the labor market is said to affect the choices of minorities in such a way that self-employment becomes a feasible alternative. Labor market disadvantages can be thus treated as resulting from a mismatch between skills (or recognition of skills) of minorities and their labor market opportunities.

Migration studies complement economic considerations with greater regard for culture-related nuances, such as issues connected to labor market integration and resulting immigrant incorporation strategies. Light (2004) provides a typology of disadvantages immigrants may face in the context of labor market activity (see Table 2.1). It follows that if labor market discrimination is in place, but one does have access to some resources, self-employment may be the best alternative.

The claim that labor market discrimination affects people's choice to become self-employed has been termed the "discrimination (or disadvantage) hypothesis". Though in the field of economics this strand of literature considers minority populations in general, here the focus will be specifically on immigrants. Over the past three decades the discrimination hypothesis received a number of formal conceptualizations and empirical

Table 2.1: Immigrant disadvantages

		resource disadvantage	
		yes	no
labor market disadvantage	yes	immigrants have low productivity and for the productivity they demonstrate they are not adequately rewarded, possibility of relying solely on the informal economy or experiencing long-term unemployment	immigrants can not obtain a wage which reflects their productivity; self-employment may be a more rewarding or the only possible source of income, if resources are sufficient it may be pursued in the formal ethnic economy or even in the open market
	no	due to low resources (human, social, cultural capital) immigrants have low productivity and therefore receive low wages; very limited possibilities of occupational mobility or pursuing self-employment in the informal economy where limited resources are sufficient	*no disadvantage*

Source: Author's own elaboration based on Light (2004).

verifications. The following three types of discrimination have been recognized: 1) employer discrimination (e.g. Moore, 1983; Clark and Drinkwater, 2000) 2) capital market discrimination (e.g. Coate and Tennyson, 1992) and 3) consumer discrimination (e.g. Borjas and Bronars, 1989). The first of these explains under which conditions would immigrants choose self-employment over wage-employment. The second concept explores access to capital as the key ingredient of entrepreneurial activity and examines how borrowing constraints affect the incentives and potential for the development of immigrant entrepreneurial ventures. The last concept is useful for explaining how consumer preferences with respect to providers of goods and services may affect the returns and thus also the numbers of immigrant businessmen.

Employer discrimination may be exerted in two forms (Parker, 2006). Either by blocking a group's access to the labor market in general, or by restricting their opportunities to low-paid jobs. Such limitations could result in choosing self-employment as a substitute strategy. Given the theoretical assumptions of formal models of employer discrimination, e.g. identical distributions of entrepreneurial abilities in the migrant and non-migrant groups and that business profits are an increasing function of these abilities, the explanatory power of this concept is not enough to answer the question of why some immigrant groups have higher and others have lower rates of self-employment when compared to the natives, though.

The second type of discrimination, discrimination in capital markets may have a direct effect not only on the choice between wage-employment and self-employment, but also on the survival rate of ethnic businesses. Not having access to bank loans or being able to borrow relatively small amounts of money and at high interest rates may heavily impede business set-up and development. It has been shown, though, that even here the discrimination performed by lenders does not necessarily affect all immigrant groups equally (Parker, 2006). One of the possible explanations arises from the specificity and clustering of immigrant businesses mostly around only a few industries which have above-average failure rates (ref.: Bonacich's (1973) middleman minority theory and Werbner's (2001) insights into the enclave economy hypothesis for discussion on immigrants' sectoral specialization). When the relatively low collateral of immigrants, as compared to the possibilities of the native population, is added to the picture, it may look as if banks discriminate immigrants as such, while in fact they "discriminate" certain types of (risky) businesses (which happen to be run mostly by immigrants), a process referred to as statistical discrimination (Arrow, 1998). Coate and Tennyson (1992) develop a theoretical model where they postulate that credit market discrimination actually occurs due to initial labor market discrimination.

The idea of consumer discrimination and its applicability to immigrant self-employment has been analyzed by Borjas and Bronars (1989), among others. The authors propose a model in which incomes and rates of self-employment among minorities are explored based on the hypothesis that consumers dislike buying goods and services from minority businessmen (see: Becker, 1971). The equilibrium outcomes of both the consumer and producer choices imply that: 1) the average income of self-employed black entrepreneurs is lower than the average income of white entrepreneurs and that 2) the gains from self-employment for able black entrepreneurs are smaller than the gains from self-employment for able white entrepreneurs. The two conclusions imply that minority entrepreneurs not only have lower incentives to become self-employed, but that they are also negatively selected into self-employment with a greater probability than the majority population.

Employer discrimination

Employer discrimination may be visible in restricting a minority's access to the labor market to simple or undersupplied occupations or to capping immigrants' earnings. In either case the effect observed on the aggregate level would be the same, lower wages of immigrants in comparison to the natives.

The immigrants' labor market restraint may apply to the whole market, making the members of minority groups unemployable, or only to some industries or occupations, e.g. those which are associated with more prestigious jobs. In either case self-employment may be an escape strategy. This concept has been also termed "blocked mobility" or "bleak mobility" (Mata and Pendakur, 1999). The former appeared in the work of Waldinger et al. (1985), Beaujot et al. (1994), or more recently by Raijman and Tienda (2000), among others. The notion refers to a situation in which immigrants pursue self-employment

due to 'glass ceilings' as far as occupational upward mobility is concerned. Because of unrecognized qualifications or discrimination based on ethnic prejudice they experience a mismatch between their skills and labor market opportunities made available to them. In their study Raijman and Tienda (2000) find that the "blocked mobility hypothesis" obtains for Koreans, who view self-employment as the "price of immigration to be paid by the first generation" (Raijman and Tienda, 2000, p. 701). It is hypothesized, though, that as time passes immigrants (and with greater probability their children) may obtain recognition for their skills or acquire skills by the standards of the receiving society, become proficient in the local language and, as a consequence, be able to eventually break out of the affliction of blocked mobility.

As far as employer discrimination is concerned, its effect on labor market choices of immigrants has been tackled formally. Decreasing immigrants' wages relative to those of non-minority workers with the same productivity potentially leaves the ratio of minority business profits relative to native business profits higher then their ratio of wages. By means of employing this comparison with the natives into the decision-making process minority members may find self-employment a relatively more attractive strategy. Regardless of what the nominal earnings in either type of activity are. The latter process is hypothesized to continue only to such a point when the crowding of migrants into entrepreneurship pushes down the average profits low enough, that

$$w_M/w_N = \pi_M/\pi_N$$

where w denotes wages, π denotes profits from entrepreneurship and indices M and N indicate minority and non-minority members, respectively.

Consequently, assuming identical distributions of entrepreneurial abilities in the minority and non-minority groups and that business profits are an increasing function of these abilities, then for the marginal entrepreneur (who is indifferent between wage-employment and self-employment) with ability $\tilde{x}$

$$w_N = \pi(\tilde{x}_N)$$
$$w_M = \pi(\tilde{x}_M)$$

and since, by definition, due to employer discrimination

$$w_N > w_M$$

then

$$\pi(\tilde{x}_N) > \pi(\tilde{x}_M)$$

what in the words of Parker (2006) expresses the notion that immigrant (minority) entrepreneurs are less able and gain smaller returns to their enterprises than majority businessmen. Nevertheless, given the assumption of equal distributions of skills in the two

populations, the lower skills required for the minority group members to switch into entrepreneurship imply a proportionally larger number of minority entrepreneurs, relative to the majority population.

As Parker (2006) observes, though, this is not true for all minority groups. Some immigrant groups have higher while other groups have lower rates than the native population on the same market. Thus it remains a question why would, according to this approach, employers discriminate against some groups, but not other ones. A possible explanation, as proposed in this thesis, lies in the fact that the theoretical model assumes a very simple mechanism of differentiating between the inner and outer group. One can either be a native or an immigrant. It is implicitly assumed that employers are natives and potential workers are either natives or minority members. Had the employers been minority members, i.e. immigrants entrepreneurs would the hypothesized relationships hold? This notion will be recalled in the empirical analysis related to immigrant labor market discrimination and its effect on immigrant self-employment. The possible preference of minority (immigrant) employers for minority (immigrant) employees will also be exploited in the empirical analysis of how ethnic concentration ratios affect ethnic business profitability.

Consumer discrimination

Similarly to employers, also customers may have preference for providing business to some, but not to other firms. Borjas and Bronars (1989) describe a model in which incomes and rates of self-employment among minorities are explored based on the hypothesis that consumers dislike buying goods and services from minority businessmen (see: Becker, 1971).

The set up of the model assumes two types of sellers - black (b) and white (w) who sell a homogeneous good. Black sellers constitute a minority of the population, the fraction of them being $\theta < 0.5$. The number of black buyers in the population is the same. It is further assumed that white buyers discriminate against black sellers, but that black buyers have no preference with respect to the race of the seller. All consumers are utility maximizers, do not discount, live infinitely and are risk-neutral.

Buyers contact sellers with cost C and prior to bearing that cost they do not know what is the race of the seller and what price he or she charges for the good. The maximum price buyers are willing to pay for a good purchased from a "preferable" seller is R. If white consumers were to buy from an "not preferable" (black) seller, they would be willing to pay maximum $R(1-d)$, where d is a discrimination coefficient. Thus the reservation price $P^*(i,j)$ offered by seller i to consumer j at which the consumer would be indifferent with respect to buying from that seller of continuing the search is given by

$$R - D(i,j)P^*(i,j) = max[0, -C + EV(P,i,j)]$$

where

$$D(i,j) = \begin{cases} 1/1-d & \text{for} \quad i=b \quad \text{and} \quad j=w \\ 1 & \text{otherwise} \end{cases}$$

and where $EV(P, i, j)$ is the expected value of the price offer given by seller i to consumer j with $i, j \in \{b, w\}$.

Sellers, on the other hand, devote their time either to production or selling. They choose the hours worked per day (H) and the fraction of consumers they serve with: $\tau = 1$ if they sell to all, $\tau = \theta$ if they sell only to blacks and $\tau = 1 - \theta$ if they sell only to whites. Furthermore, in the model proposed by Borjas and Bronars (1989), sellers differ in their ability to produce the market good and can be either high-skilled or low-skilled and the skills are equally distributed among the w and b sellers. The sellers maximize utility

$$U = \frac{\alpha\tau\beta}{\alpha\tau+\beta}P(\tau)H - \frac{H^{\delta}}{\delta}$$

where α is the number of transactions any seller (be it high- or low-ability) can complete per unit of time, τ is the fraction of contacts that result in a sale, β is the ability (productivity) of the seller, $P(\tau)$ is the price asked by the seller dependent on his segregation strategy, H is the number of hours worked and δ is a parameter greater then 1, what makes the relationship between utility from income and disutility from hours worked nonlinear.

From the above set up and by determining the equilibrium outcomes of both the consumers and sellers choice, Borjas and Bronars (1989) conclude that: 1) the average income of self-employed *black* entrepreneurs is lower than the average income of *white* entrepreneurs and that 2) the gains from self-employment for able *black* entrepreneurs are smaller than the gains from self-employment for able *white* entrepreneurs. The two conclusions imply that minority entrepreneurs not only have lower incentives to become self-employed, but that they are also negatively selected into self-employment with a greater probability than the majority population. Here, similarly to the theoretical model of employer discrimination two assumptions would have to be made in order to extend the theory to immigrant discrimination. First, it would have to be assumed that a minority defined by race may just as well be a minority defined by immigration status. On one hand this is a fair assumption as customers, contrary to .e.g employers, have no possibility of checking an entrepreneur's immigration status. On the other hand, this is faulty, though. Due to e.g. the capability for self-exploitation, immigrants may prefer to work long hours in order to achieve their financial migration goal sooner and return home. Thus they may establish their businesses in (sub)sectors where the majority doe snot find it profitable to work, leaving the local consumers no choice and disabling discrimination all together. Second, among the consumers we may just as well see minority members, rather than the majority population. In that case we should consider that they could possibly discriminate against majority businesses.

Credit market discrimination

The final market in the context of which minority discrimination has been considered in relation to entrepreneurship is the credit market, more often than not an unavoidable economic arena for business ventures. Coate and Tennyson (1992) develop a theoretical

model, which provides the intuition of how credit market discrimination is in fact a derivative of initial employer discrimination, i.e. how can employer discrimination "spill" into other segments of the market and affect both the odds and the gains to self-employment. In their set up Coate and Tennyson (1992) recognize two distinct groups, which could be different either in terms of race, ethnicity, gender or any other exogenous characteristic. The labor market offers two jobs for two types of workers - "skilled" and "unskilled". The marginal product of either skilled or unskilled workers in the unskilled occupation is w_u. In the skilled occupation the marginal product of skilled workers is w_s, while that of unskilled workers is 0.

Marginal productivity of a skilled worker is strictly larger than the productivity of an unskilled worker. The model further assumes that the fractions of skilled and unskilled workers in the two groups (let us continue with minority (M) and non-minority (N) workers) is the same. Yet due to employers' discrimination practices against the minority group (restricting access to skilled jobs), the probability that a skilled minority member obtains a skilled job is $\pi < 1$. The probability that a majority skilled worker obtains a skilled position is 1. Unskilled workers, be it minority or majority, are always assigned to the matching unskilled job. Workers are further assigned an exogenous probability of success in case of undertaking entrepreneurial activity $p \in [0, 1]$. For unskilled workers this probability is equal to 0. For skilled workers, minority and non-minority, the allocation of entrepreneurial ability follows an identical distribution function for both groups *G(p)* with the probability density function *g(p)* continuous and positive on the range of [0,1]. It is finally assumed that each worker has to borrow $K > 0$ capital at interest rate r in order to start a business and the returns from a successful venture are R, and 0 in case of failure.

Given this set up a risk-neutral, skilled worker will choose to become self-employed if, given a certain probability of success (p), the returns from business (R) net of the costs of borrowing capital ($K + K * r$) are larger than the expected wage w. Thus, the marginal borrower for whom the expected return from wage-employment and self-employment would be the same can be described as facing the following equality

$$p^*(r, w) = w/[(R - (1 + r)K)].$$

A credit market equilibrium requires that expected profits from lending to either minority or majority entrepreneurs be equal to 0. Thus, given the risk-free interest rate q and the expected probability of success $E[p|p \geq p^*(r_i, w_i)]$, the interest rate pair (r_M^*, r_N^*) would be an equilibrium if

$$(1 + r_i^*)E[p|p \geq p^*(r_i^*, w_i)] = 1 + \rho, \qquad i \in \{N, M\}$$

In case of (skilled) minority members, who, due to employer discrimination, have to expect a lower wage in the labor market the profitability of business, which would create a sufficient incentive for them to switch to self-employment is lower, than for the non-minority population. Following the logic of the model as it has been set up by Coate and Tennyson

(1992) the expected probability of success (which depends positively on the expected wage) for minority entrepreneurs will be lower than for the non-minority entrepreneurs. Thus the interest rate for immigrant borrowers has to be higher in order to compensate the lenders for increased risk of the investment. This mechanisms has been termed "credit market discrimination". It has been know to be purely "statistical discrimination", though (see: Arrow, 1998), i.e. a case of discrimination in which members of a certain group are not discriminated against because of their membership in the group *per se*, but because this group is characterized by a specific set of attributes.

The higher interest rate resulting from the labor market and credit market discrimination further implies lower returns from self-employment for a minority entrepreneur with the same probability of entrepreneurial success as a majority entrepreneur. Nevertheless, the expected return from self-employment relative to wage-employment is higher for the minority than for the majority businessmen, thus creating a stronger incentive to become self-employed among the former.

In a further development of the model Coate and Tennyson (1992) allow the workers to invest (I) or not invest (NI) in acquiring additional skills, thus partially endogenizing the skill variable. Under such a set up it is shown that lenders will differently approach the groups who invested in their human capital and those who did not. Under the assumption that minority and majority members invest optimally in their skills with the net utility gain from investing being $\Delta(r_i^*, w_i)$, the interest rates (r_N^*, r_M^*) will result in an equilibrium if the following equalities are met

$$(1 + r_i^*)E_I[p|p \geq p^*(r_i^*, w_i)] = 1 + \rho \quad and \quad \Delta(r_i^*, w_i) \geq 0, \quad i \in \{N, M\}$$

or

$$(1 + r_i^*)E_{NI}[\text{p—p}\geq p^*(r_i^*, w_i)] = 1 + \rho \quad and \quad \Delta(r_i^*, w_i) < 0, \quad i \in \{N, M\}.$$

In such a case the relations of interest rates offered to minority and majority entrepreneurs who invested in skills and who did not would be as follows

$$r_{N(I)} < r_{M(I)} < r_{N(NI)} < r_{M(NI)}.$$

Yet, due to the fact that the investment in skills is assumed to be unobservable to the lender, the interest rate for either the minority or majority has to be the same. Under this assumption incentives to become self-employed are further explored. Coate and Tennyson (1992) observe that under specific conditions (see lemma in: Coate and Tennyson, 1992, p. 283) one of which is that it is possible that the minority (the discriminated group) with the same entrepreneurial abilities as the majority (the non-discriminated group) will have less incentive to enter self-employment. This is to say that there exist such interest rates with which not only a smaller proportion of the minority would end up skilled (relative to the majority), but out of those who are skilled fewer would become self-employed.

Finally, Coate and Tennyson (1992) look into a case when investment decisions of individuals are observable to the lender and thus the net gain from investing takes the form $\Delta(r_i(I), r_i(NI), w_i), i \in \{N, M\}$. Yet, if a minority worker was to invest in his or her

skills but the stimulus of a reduced interest rate would not compensate for the deterrent effect of relatively low wages, then it would still be possible that minority members have less incentive to enter self-employment, as compared to the majority of the population.

In this manner Coate and Tennyson (1992) show how discrimination in capital markets may have a direct effect not only on the choice between wage-employment and self-employment, but also on the returns to entrepreneurial activity and the survival rate of minority businesses. Not having access to bank loans or being able to borrow relatively small amounts of money or at high interest rates may heavily impede business set-up and development. It has been shown, though, that even here the discrimination performed by banks does not necessarily affect all immigrant groups equally, Parker (2006). In the case of immigrants as a minority population, one of the possible explanations of this fact arises from the specificity and clustering of immigrant businesses mostly around only a few industries which have above-average failure rates. This is why it may look as if banks discriminate immigrants as such, while in fact they "discriminate" certain types (risky) of businesses, which happen to be run mostly by immigrants. Here, juts as in the other two types discrimination discussed heretofore the conclusion is valid as long as the assumption that the one who provides the credit is from the "other" group than the one who applies for the credit. This assumption may be easily violated if immigrants (knowing that their borrowing is constrained on the receiving market) borrow from within their ethnic groups, i.e. relay on the support of their immigrant network. This feature will underlie the reasoning developed in the empirical analysis of ethnic concentration ratios and their effect on ethnic entrepreneurship.

Moderators of discrimination

Labor market discrimination of immigrants can be moderated by means of several mechanisms. Among them we find the duality of labor markets and the highly related notion of relative deprivation. The former has been developed by Doeringer and Piore (1971) and Piore (1979). Despite the fact that the concept was not tailored to fit the issue of immigration, it has been widely used to explain immigrants' labor market behavior. Doeringer and Piore (1971) recognize the duality of labor markets in industrialized economies, namely that there are two distinct sectors, the primary and the secondary. The primary sector is characterized by relatively stable employment opportunities and well-payed jobs in which human capital is effectively employed. The secondary sector experiences opposite features and comprises 3-D jobs, i.e. dirty (dull), dangerous and demeaning. Such a market structure and especially demand for labor in the secondary sector is what creates a pull-factor for migrants. The unwillingness of native workers to undertake jobs which are connected to low social status leaves a number of positions open to potential newcomers. Labor market duality also theoretically enabled migration in presence of unemployment, as the segments of the markets on which there was unemployment and on which there was demand for labor need not be the same one.

The duality of the labor market affects discrimination in the sense that it hypothesizes that there are markets, e.g. related to specific industries, where not one's skills are critical, but rather the willingness to perform menial work. A highly skilled worker would not be rewarded according to his level of productivity considered in absolute terms. On the secondary labor market his or her chances of demonstrating and making use of this productivity are limited. Thus on top of wages being differentiated based on measures unrelated to productivity, e.g. race or gender, they are also differentiated based on the unequal access of some groups to certain segments of the labor market, so called occupational discrimination. Labor market duality may thus inflate the difference in earnings between the immigrants and the natives, providing an even stronger push factor for self-employment, than wage discrimination alone.

The second mechanism which may potentially moderate the effect of discrimination by utilizing the notion of labor market duality is relative deprivation (Stark and Bloom, 1985), considered within the framework of the New Economics of Labor Migration. Based on a model in which utility is determined not only by one's income but also by the relative value of that income as compared to the incomes of others, Stark and Bloom proposed that people may migrate in order to change their group of reference. This hypothesis was developed upon the premise that "...people engage regularly in interpersonal income comparisons [which] generate psychic costs or benefits..." (Stark and Bloom, 1985, p. 173). People might prefer to be the richest among the poorest abroad, than the poorest among the rich at home. By working in the secondary sector, where all immigrants, who are assumedly the comparison group for other migrants, are more or less equally worse off than the natives working in the primary sector, immigrants decrease the psychic costs of being different, thereby increasing their utility.

Given the notion of relative deprivation and the duality of labor markets, discrimination may have a deterring effect on the propensity to become self-employed. As it has been described in the previous section, employer discrimination induces migrants' crowding-in into self-employment as long as

$$w_M/w_N < \pi_M/\pi_N$$

If the comparison group, following the logic of Stark and Bloom (1985) are not natives, but other immigrants working in the secondary sector, then the wage difference between immigrants and their comparison group might be smaller than that between migrants and natives. In this sense the idea of relative deprivation may deflate the wage differences considered as a decision-making factor for immigrants, weakening the incentives to choose self-employment over wage-employment.

Summary

Summarizing, labor market discrimination is a factor which may encourage people to undertake self-employment. Discrimination of immigrants takes the form of a wage disadvantage and an occupational disadvantage. The latter is related to the duality of the labor

market and the fact that immigrants are more often than not employed in the secondary sector. Thus, on one hand, the duality of labor markets amplifies the effects discrimination. On the other hand, working on similar terms as other immigrants decreases one's level of relative deprivation what potentially increases one's utility from migration and wage-employment. Whether the utility from wage-employment is large enough to keep people from switching to self-employment is a debatable issue. It seems relevant to note who constitutes the comparison group in relation to which one evaluates his or her earnings.

2.3.2 Ethnic economies

So far we have discussed discrimination as a consequential circumstance which affects the propensity to become self-employed among members of immigrant groups. The second strand of literature which provides an important context for the analysis of ethnic entrepreneurship is related to the development of ethnic economies. This stream of thought put more emphasis on the internal economic structures and modes of operation of minority communities, rather than on the relations of ethnic minority entrepreneurs with the receiving society. Such a perspective gains importance as minorities become more and more numerous, being able to become self-sufficient communities within the receiving markets and societies.

The enclave economy hypothesis was developed by Wilson and Portes (1980) based on an analysis of longitudinal data on Cuban immigrants in the US Wilson and Portes (1980) tested the hypothesis that there exists a third alternative to the postulated primary and secondary labor markets (see: Doeringer and Piore, 1971; Piore, 1979). By analyzing the labor market incorporation of Cuban immigrants they found a significant difference between migrants who worked in the peripheral economy (companies in sectors with relatively low average wages, relatively small average employment and without internal promotional ladders) and those who worked for Cuban entrepreneurs. Cubans working for Cuban employers (what was the identification of functioning within an enclave economy) were found to experience significant returns to their human capital, similarly to workers within the primary labor market. In the open, secondary labor market such returns were said to have been absent because immigrants did not have an opportunity (or need) to take advantage of their culture-specific human capital. Wilson and Portes (1980) claimed that, from the immigrant entrepreneurs' point of view, hiring labor from within the same immigrant community resulted in opportunities for expansion due to privileged access to markets and labor, or ethnic solidarity and obligation of reciprocity (see also: Bonacich, 1973). The two conditions which were said to have been necessary for the development of immigrant enclaves were: 1) access to sufficient start-up capital (either through immigrant linkages or by connections with the home country) and presence of entrepreneurial skills among some people belonging to the immigrant population, and 2) the renewal of the labor force within the enclave through immigration. The main benefits of operating a business within an ethnic enclave were thus: access to labor, access to capital, and access to a pool of consumers.

The enclave as a source of labor

As far as the role of ethnic enclaves in proving cheap and reliable labor for ethnic entrepreneurs is concerned, Sanders and Nee (1987) challenged the enclave economy hypothesis as proposed by Wilson and Portes (1980). Based on an analysis of earnings of Cuban and Chinese immigrants they found that, indeed, functioning within an enclave economy may be beneficial for immigrant entrepreneurs, but that it is not necessarily the case for their co-ethnic employees. After adapting Wilson and Portes' methodology Sanders and Nee (1987) re-examine the Cuban population of Miami and Hialeah, Florida and the Chinese enclave in San Francisco, California and note that: 1) immigrants socioeconomic achievement is negatively related to their spatial concentration in ethnic enclaves and that 2) immigrants' socioeconomic achievement is positively related to their level of assimilation. These findings lead to concluding that an enclave creates beneficial conditions for newcomers' incorporation into the labor market and for early stage entrepreneurs, but that from a longer time perspective there is also a cost to the segregation from the mainstream economy. Sanders and Nee (1987) find that the disadvantages related functioning in an enclave, as opposed to functioning in an area with a smaller concentration of one's co-ethnics, apply to the wage-employed. They did not find any significant difference between operating in and outside of the enclave for the self-employed.

Sanders and Nee's (1987) approach has been criticized by Portes and Jensen (1989) who suggest that operationalizing enclaves as places where people reside, rather than where they work, as it was done by Sanders and Nee (1987), is conceptually wrong. They reexamine the Cuban population in Miami and compare individuals' place of work and place where their household was enumerated. They find that indeed the two cannot be used interchangeably, as 1) the proportion of Cubans for whom these two locations do not overlap is sizable (40%) and 2) the populations for which the two locations overlap significantly differ in their characteristics from those for which these two locations are not the same. In subsequent analyses of earnings of Cubans living in ethnic enclaves in Florida (at this time they also account for immigrants from the "Mariel Boatlift", 1980), Portes and Jensen (1989) look at the estimators of dummy variables indicating whether one is self-employed or wage-employed. Comparing the estimator of the self-employment dummy across models run on a sample of men employed in Miami (the ethnic enclave) and men employed outside of Miami, they conclude that being an entrepreneur is positively related to one's earnings if he or she works in the enclave and negatively related if one operates a business outside of the enclave. Thus, they reject the hypothesis formulated based on Sanders and Nee's (1987) findings, that there is no positive effect of enclave entrepreneurship on earnings. They also find support for a rather positive role of the ethnic enclave in providing opportunities of socioeconomic mobility for immigrant workers and providing access to resources and capital for immigrant entrepreneurs. The issue with the empirical strategy applied by Portes and Jensen and Sanders and Nee is that the effect of being self-employed on earnings is captured by means of a dummy variable. Such an approach assumes that being self-employed is related to the one's earnings independently

to all the other explanatory variables. From the discussion on the determinants of self-employment it follows, though, that this must not be the case. The demographic variables affecting ones productivity (earnings) are also those which significantly affect the process of selection into self-employment as a form of labor market activity.

Zhou and Logan (1989) conduct further studies on the enclave economy by exploring the case of the Chinese in New York City. In order to provide robust results, irrespective of what we consider to be enclave, they approach three possible meanings: 1) that of a place of living, 2) that of a place of work and 3) that of an industry. In the latter conceptualization (not considered in the previous literature) they identify enclave industries as those, where the Chinese immigrants are over-represented. They manage to reconcile the findings of Wilson and Portes (1980) with those of Sanders and Nee (1987). On one hand they find support for the positive view of the enclaves role, specifically in relation to the possibility of upward mobility of immigrants via enhanced opportunities for self-employment. On the other hand, they do not find advantages in terms workers earnings or the entrepreneurs returns to self-employment when individuals within and outside of the enclave economy were compared.

The debate (Jiobu, 1988; Portes and Jensen, 1989; Model, 1992; Sanders and Nee, 1992) on the nature or ethnic enclaves leads Portes and Jensen (1992) to a revision of their initial proposition. They eventually describe the role of the enclave as a local economy which does offer employment comparable in its returns to the mainstream economy, but that this feature holds rather for recent immigrants and to those who have limited language proficiency. The refinements and reconceptualizations of the enclave economy hypothesis proposed over the course of the first decade after the idea was introduced by Wilson and Portes in 1980, suffered from similar drawbacks of the empirical analyses they were based on. They all concerned de facto case studies of certain minorities living in specific locations. The inquiries into the costs and benefits generated by ethnic enclaves predefined them as localities where there are very high concentrations of given ethnic groups. The extent to which the changes in the level of this concentration affect the role of the enclave as limiting or boosting opportunity remained unexplored.

Referring to the various approaches to the ethnic enclave economy hypothesis Light et al. (1994) come to a conclusion that what has developed in the literature so far is "conceptual anarchy" (Light et al., 1994, p. 69). In terms of definitional issues Light et al. (1994) assert that the interchangeable use of the terms ethnic economy and ethnic enclave economy is conceptually wrong. Their reasoning stresses that the ethnic economy is a concept derived from the middleman minority theory (Bonacich, 1973), while the notion of an enclave arouse from the theory of labor market segmentation. Aside from clarifying the semantics Light et al. (1994) emphasize that analyzing relative (within-enclave vs open market) wages as a measure based on which the enclave economy hypothesis is tested is misleading. The related assumption being that the people who earn wages outnumber those whose activity is not regulated by any employment contract. As Light et al. (1994) show, based on census data on ethnic employers and employees, in case of ethnic minori-

ties the opposite is true - the employers outnumber their employees, as many people are actually self-employed and those who can afford to hire workers do that on a very small scale. Thus, as Light et al. (1994) conclude, the relative wages may be a very deceiving measure of the welfare of the participants of the ethnic economy. Yet, as the authors further state, even if the wages of employees in the ethnic economy were actually lower than what they could earn in the open market, this does not indicate that an ethnic (enclave) economy creates a mobility trap, as it has been postulated by Portes and Wilson's opponents. In fact three issues should be considered: 1) that some employees in the ethnic economy may not want or may not be able to work full time, thus earning lower wages, but also voluntarily working fewer hours, 2) the ethnic economy could also be considered as a school for entrepreneurs, bringing the minority long-term benefits rather than advantages instantly visible in income data, and 3) comparing wages in the ethnic economy to those in the open market is fair only if the open market offers suitable jobs to everyone looking for them (in fact, in the open market one could be underemployed or even unemployed and so compared to earning nothing even a low wage in the ethnic economy is supreme). Thus Light at al. summarize the relative wages debate by claiming that relative earnings in the ethnic economy determine only the extent to which it is beneficial to work there, not whether the ethnic enclave creates a beneficial environment in general.

The enclave as a source of financial capital and consumers

The idea of how the role of enclaves in providing employment and hiring opportunities corresponds to the fact that both the workers and employers are at the same time potential consumers is a feature under-explored in the mainstream ethnic enclave debate. The notion of co-ethnics being potential creditors has also been signaled, but not tackled conceptually. In line with Waldinger et al.'s (1985) previous inputs to the disadvantage theory of immigrant self-employment, in 1990 the authors suggested a conceptual framework in which they combine the idea of immigrants limited opportunities (e.g. due to discrimination) with their possibility of mobilizing ethnic resources (e.g. within ethnic enclaves). Due to the mutual influences of its various components, the framework has been named "interactive". The framework is created based on an analysis of existing theoretical concepts and statistics of ethnic entrepreneurship.

The opportunity structures the authors define include historically shaped circumstances which enable (and constrain) ethnic entrepreneurship - market conditions and access to ownership possibilities. They recognize that the types of industries, where immigrants are most likely to be able to set up the enterprises constitute specific niches: 1) where mass production technology does not apply and where mass distribution is unnecessary, 2) where there are low economies of scale, 3) where there is instability and uncertainty, and 4) where ethnic goods are in demand.

Apart from the somewhat external factors creating opportunity structures Waldinger et al. (1990) suggest that the immigrant group's own characteristics also affect the rates of entrepreneurship. These characteristics are divided into predisposing factors and possibil-

ities of resource mobilization. Among the former the authors recognize blocked mobility, selective migration and migrants aspiration levels. The latter category encompasses ties with co-ethnics, extent of social networks and government policies.

Ways in which ethnic entrepreneurs take advantage of the opportunity structures, given their group characteristics Waldinger et al. (1990) label ethnic strategies. These strategies are aimed at tackling the seven most common problems: 1) obtaining information, 2) generating start-up or development capital, 3) acquiring necessary entrepreneurial skills, 4) labor recruitment, 5) establishing and developing relationships with customers and suppliers, 6) dealing with competition and 7) responding to political attacks. The possibilities and eventual ways of addressing these issues emerge from the constraints and enabling mechanisms embedded in the market structures and ownership possibilities as well as from the specificity of predisposing factors and the potential of resource mobilization.

Within this framework Waldinger et al. (1990) look for explanations of why self-employment rates may differ across ethnic groups. They define three categories by which the immigrant communities differ, and which affect their functioning in the destination countries: 1) pre-migration characteristics (mainly level of human capital), 2) the circumstances of migration and the ways they evolve (e.g. into temporary vs permanent migration) and 3) post-migration characteristics (especially the position of the ethnic group in the host economy). Though Waldinger et al.'s (1990) framework was not formally tested, it provides valuable guidance for empirical analysis.

Social and cultural capital as a resource

As social and cultural capital have been recognized as an ethnic resource, more and more concepts have relied on this premise. Among the most influential we may find the idea of mixed embeddedness or the frameworks of modes of incorporation.

Kloosterman and Rath (2001) develop the concept of mixed embeddedness with an intent to provide a solid theoretical framework suitable for international comparisons of immigrant self-employment. Mixed embeddedness implies that immigrants are not only embedded in immigrant networks and their ethnicity, but also in the "socioeconomic and politico-institutional environment of the country of settlement" (Kloosterman and Rath, 2001, p. 2). In their in-depth inquiry into what has so far been labeled generally as opportunity structures (see: Waldinger et al., 1990) Kloosterman and Rath firstly recognize, that immigrant communities not only differ from the majority in "cultural" terms, but also as far as other, more tangible, forms of capital are concerned - financial, human, social. Thus, they are initially dependent on different segments of the opportunity structure than the native population. The second aspect of opportunity structures as understood heretofore, which Kloosterman and Rath (2001) question, is their static character. Migrants may, by their mere presence, change opportunity structures. This dynamic aspect has been under-explored as most analyses of immigrant self-employment are case studies of specific ethnicities in specific locations at a specific point in time.

In their furthering of how opportunity structures should be viewed, Kloosterman and Rath (2001) recognize two crucial dimensions - the accessibility and growth potential of markets. In effect Kloosterman (2004) proposes a two-dimensional typology of their interrelations. The typology allows to systematize various markets in terms of their attractiveness and accessibility for immigrant businesses by classifying them as stagnant or expansive on one hand, and low threshold or high threshold (as far as human capital requirements are concerned) on the other hand. Ethnic markets are according to Kloosterman (2004) able to emerge in all but one categories of this typology and thus should be treated as special cases of protected markets within the economy (see also: Wilson and Portes, 1980). The only case which Kloosterman (2004) finds unattractive for immigrant businesses is that of stagnant and human capital demanding markets. This notion is in line with the early observation of Wilson and Martin (1982), who conducted an input-output analysis for Cuban and black businesses in Florida. They found that in the construction, manufacturing and "other" industries Cuban firms had greater potential to buy most of their required inputs from other Cuban firms in the area. This was not the case for black firms, which would have to be supplied by non-black firms. For trade, transportation and services the structure was similar for both populations. They concluded that the Cuban population is more integrated business-wise than the black population. This enabled explaining why some minority enclaves fare better then others. Wilson and Martin (1982) estimated that the better integrated Cuban economy could generate an additional 50% of spending within the community once an initial demand was injected. This last conclusion is based on a strong assumption, though, namely that entrepreneurs would prefer intra-community business to market exchanges with other groups. Given the theory behind the functioning of ethnic networks this is not an unreasonable expectation, yet in case of entrepreneurship this may be highly industry-specific (Nestorowicz, 2008). Given the specificity of immigrants as opposed to native minorities, the comparability of the two groups in Wilson and Martin's (1982) study also leaves place for doubt.

The contexts and opportunity structures under which immigrant self-employment becomes a feasible labor market strategy have been further elaborated upon in concepts which can be classified as frameworks of modes of immigrant incorporation. A major contribution to this stream of thought was that by Portes and Rumbaut who developed a typology of immigrant incorporation with ethnic entrepreneurship as one of its features. According to them what is crucial to recognize when trying to understand the variations in immigrants' labor market performance are: 1) immigrants' resource endowments in the form of educational attainment, skills, motivations, aspirations, professional experience, financial, social and cultural capital etc., 2) their status of entry and residence (legal, irregular, undocumented), 3) the conditions under which they left their countries of origin, and 4) the contexts of reception. The latter have been decomposed into governmental, labor market and ethnic community aspects.

Waldinger (1987, 1996) analyzes a specific case of immigrant incorporation, namely that which took place in the context of industrial change in the second half of the twen-

tieth century in New York City. Waldinger conceptualizes the process of how immigrants manage to undertake self-employment due to succession of entrepreneurial openings (Waldinger, 1987, 1996) and labels it the game of "ethnic musical chairs" (Waldinger, 1996, p. 257). In Waldinger's view the industrial transformations taking place in the 70s in the New York City area lead to changes in the composition of local industrial and labor markets such that the numbers of whites declined "set[ing] in motion a vacancy chain, allowing non-whites to move up the job hierarchy as replacements for whites" (Waldinger, 1987, p. 370). Waldinger observes that the small business segment of the market went trough a similar succession process (Waldinger, 1996). In the above-mentioned circumstances self-employment became a predominantly immigrant activity, with rates exceeding those of African-Americans and, as duration of stay increased, also exceeding those of the white population. He postulates three main reasons for such a state of things:

- increasing immigrant populations, creating their own ethnic demand which co-ethnics were best suited to serve (immigration being also a critical element of the development of ethnic enclaves),
- opportunities for succession in small business industries, which could not benefit from mass production and/or mass distribution practices and which in the new economic environment seemed mildly profitable for communities which use to occupy them,
- lack of other opportunities for immigrants whose skills did not match the labor demand or who could not have their skills recognized (a feature noted in the enclave economy hypothesis, but mainly conceptualized within the disadvantage theory).

Despite Waldinger's ability to conciliate various analytical approaches within one framework, Rath (2000) criticizes such a way of thinking when applying it to immigrant entrepreneurs in Amsterdam. He recognizes four aspects of Waldingers concept which do not fit into the stories of immigrant incorporation in the Dutch case. The first is the notion of a "labor queue" (Waldinger, 1996, p. 26). The idea implies that there is a fixed hierarchy of preferences towards specific categories of the labor force. As Rath points out, the Dutch society is not as race-conscious, though, as the American society. Moreover, he recognized the socio-political dynamics of how minority statuses might change over time, thus changing the ordering of the "queue". Secondly, Rath comments on Waldingers assumption of long-term cohesion, solidarity and support within ethnic communities. Rath notices, that the social relationships within ethnic groups tend to change over time and, especially under the circumstances of harsh ethnic competition within ethnic niches (see also: Kloosterman, 2004), this may not be true. Rath's third argument relates to the importance of local institutional frameworks at the destination, which affect the opportunities for immigrant self-employment. He criticizes Waldinger for stressing the role of ethnic networks in the process of immigrant incorporation, while the notions of consumer demand, technological change and international division are put aside. Finally, Rath draws attention to the underexposed variety of the scopes and scales of institutional frameworks

which affect immigrants' opportunities, a notion which was later emphasized by Kloosterman and Rath in their 2001 paper. In a subsequent piece of critique Rath (2001) notices also the drawbacks of Waldingers definition of niches which is said to: 1) too weakly point to the voluntary character of their formation and 2) too strongly emphasize the absolute size of the self-employed and wage-employed ethnic community (what in case of Amsterdam would lead to excluding economically and culturally significant and distinct ethnic groups). He also points to the lack of differentiation between occupations and branches of trade and industry, which may be spread over different labor markets, making the distinction crucial to understand the immigrants' modes of incorporation in these specific arenas of economic exchange. Rath's comments as to the competition and industrial specificity are one of the first in this strand of literature. Rath based his arguments on indirect evidence from existing statistical data and empirical analyses. In this thesis especially industrial specificity and ethnic competition will be empirically tested as factors affecting opportunity structures of immigrants.

Summary

Summarizing, it seems that modern day "theory" of immigrant self-employment in the context of ethnic enclaves is not far from the "conceptual anarchy" described by Light et al. (1994). This seems to be the price for being very conscious about both the constraints and the opportunities related to migrants' and entrepreneurs' activities. The ethnic enclaves hypothesis considers either business expansion or cost minimization, both of which lead to relatively higher returns, fitting well into the income maximization objective of immigrants and entrepreneurs. In the absence of an explicit income advantage of being self-employed, the ethnic enclave hypothesis enables considering enclaves as an engine for one's upward mobility. Nevertheless, this may just the same be treated as the maximization of discounted lifetime income.

The numerous aspects related to the functioning of immigrants within ethnic enclaves may be grouped into features related to hiring opportunities, credit opportunities and clientele opportunities. All three relay on the social and cultural capital one shares with his or her co-ethnics. What seems to be a neglected issue, though, is that all members of a given ethnic community may just as well take advantage of these forms of capital. Such intuition begs the question of whether large concentrations of a given ethnic group do not create disadvantageous conditions for the operation of subsequent ethnic businesses. It is particularly this aspect of ethnic economies that the chapter on the role of ethnic competition and complementarity will tackle.

2.3.3 Theoretical premises for empirical analyses

The theories and concepts presented in this chapter provide ample space for posing new research questions and hypotheses. As this thesis focuses on labor market issues and labor migration, the most basic premise arising from both strands of literature is that both the self-employed and migrants are predominantly driven by relatively high returns. Testing

of whether or not this is the case will be the subject of the first empirical analysis provided in Chapter 4.

A potential affirmation of the idea that the earnings of self-employed immigrants are higher than what they could earn at home or in wage-employment would open a new field of inquiry. Immigrant entrepreneurship could be then considered as as an opportunity for those who are constrained by labor market imperfections such as wage and occupational discrimination. In this respect the notion of relative deprivation constitutes an argument for considering migrants' wage comparisons with various other groups, rather then just the majority. In presence of discrimination profitable immigrant self-employment could disentangle people from both the limitations imposed by international boundaries, as well as by culture-based prejudice encountered on the host market. The relationship between labor market discrimination and the probability of becoming self-employed will be thus the subject of the Chapter 5.

Given that the posed hypotheses are positively verified and that indeed labor market discrimination increases the concentration of immigrants in relatively profitable self-employment, it remains to be determined which mechanism may drive this positive outcome. The ethnic enclave hypothesis provides a possible answer to this question. The fact that ethnic enclaves can be a panacea to all forms of discrimination encountered by immigrant entrepreneurs is a solid theoretical basis for a claim that operating business within larger clusters of co-ethnic firms is related to increased profits. Of course it would be naive to assume that this relationship is linear. Whether immigrant enterprises actually benefit or loose from high concentrations of co-ethnic customers and competitors will be the subject of Chapter 6.

Chapter 3

The case of Puerto Rico

The United States is considered as the receiving country in all studies conducted within the framework of this thesis. The primary reason for doing so is methodological. Migrants are minority populations. Among them the self-employed also constitute a minority. In order to be able to analyze such a subpopulation quantitatively, i.e. have a sufficient number of observations of interest in a sample, the general population must be sizable.

The US is the largest immigration country in the world, with 42 million residents born abroad (United Nations Population Division, 2010). That is roughly 14% of the total population. Among such a large number of immigrants the number of the self-employed is large enough to enable empirical testing of quantitative hypotheses.

On top of the above condition that data had to enable estimations of returns to entrepreneurial activity. The US is one of the few countries which collects and provides public access to such data.

For the purpose of analyses conducted within this Puerto Rico has been chosen as the sending country. The reasons behind this decision are both methodological and conceptual. Methodologically, the benefit of considering Puerto Rico is that it provides data which is consistent with that collected in the US. This implies that just like the US, Puerto Rico also provides information on the income from self-employment. The consistency of data collected in the US and Puerto Rico arises from the fact that as of 1898, Puerto Rico is a US possession[1]. As such Puerto Rico is subject to the same schedule, scale and scope of censuses as the US.

The dependency of Puerto Rico from the US also results in the fact that Puerto Ricans have unlimited access to the US labor market. Thus, considering Puerto Rico-US migration allows to capture migratory movements which do not suffer from common conceptual drawbacks such as the immigration policy-driven selectivity of migration. The operation of political, i.e. non-market, forces in the process of migration is usually of great concern to researchers. Selective migration policies more often then not are welcoming for the "best and brightest". Immigrants are screened against their usefulness for the host economy. Such a process implies positive selection and an inevitable overestimation of

[1] In this sense US-Puerto Rico relations can be compared to those of France and Martinique, of Portugal and the Azores, of the UK and Cayman Islands, or of the Netherlands and Aruba.

the success rate of immigrant businesses, as the decisions to become self-employed are partly made for the immigrants by politicians. Those who might be subject to failure are discouraged or even precluded from starting a business venture. In this context considering Puerto Rico-US migration allows to capture the purely economic aspects of migrants' decision-making.

The last, but not least, similarity between Puerto Rico and the US is that Puerto Rico uses the US dollar as its official currency. This feature enables neglecting the exchange rate as a factor in the migration decision-making process.

On one hand, Puerto Rico is similar enough to the US to have perfectly comparable data and no political or exchange rate interference on migration, but on the other hand it is geographically and culturally distant enough to enable a clear identification of migration occurrences defined as significant changes in one's institutional setting (Eisenstadt, 1953).

Puerto Rico is an island located on the Atlantic Ocean, east of the Dominican Republic with its capital in San Juan (see Figure 3.1). Geographically Puerto Rico is further from the US than Haiti, Cuba, Jamaica or almost any location in Mexico. The distance between San Juan and Miami - the capital of the nearest US state - is around 1,600 km.

Figure 3.1: Puerto Rico - geographic location.

Source: Google Maps.

The society is predominantly Spanish-speaking, embedded in a culture which is a mix of North American and Latin, African, and Caribbean traditions (Encyclopedia Britannica, 2012). Despite the formal dependency from the US, Puerto Ricans are a self-governing political entity with a self-elected governor and bicameral parliament. They do not have the right to vote in the US presidential elections and have inconsequential representation in the US Congress. With a value of USD 16,300 Puerto Rico ranks 73rd in the world in terms of level of PPP-adjusted GDP per capita (2010 est.). The respective value for the US is USD 49,000, what gives the 11th position in the world (2011 est.; Central Intelligence Agency, 2012).

The lack of mobility restrictions on one hand, and the institutional separateness in Puerto Rican-U.S. relations on the other hand makes Puerto Rico a perfect subject to study individual migration decisions. All that is common allows to neglect the effects of selective immigration policies. The selectivity which remains stems from the socio-demographic characteristics of individuals and their responsiveness to market forces. All that contributes to a socio-economic gap between Puerto Rico and the US potentially motivates mobility and allows to recognize it as an international movement.

Despite the conceptual and methodological benefits of analyzing migration from Puerto Rico to the US, the conclusions derived from such analyses are limited tot he specific case study. In order to provide external validity of the obtained results each of the three empirical chapters devoted to testing one of the operational hypotheses will include an extension going beyond the Puerto Rican case.

Chapter 4

Gains to immigrant self-employment

4.1 Introduction

This chapter deals with a dual choice - on one hand that between migration and immobility, and on the other hand that between wage-employment and self-employment. As discussed in the previous chapter, theorists in field of migration suggest that net income gains are the predominant driver of mobility. In other words, *ceteris paribus*, one would have to earn more abroad than at home for migration to become a noteworthy option.

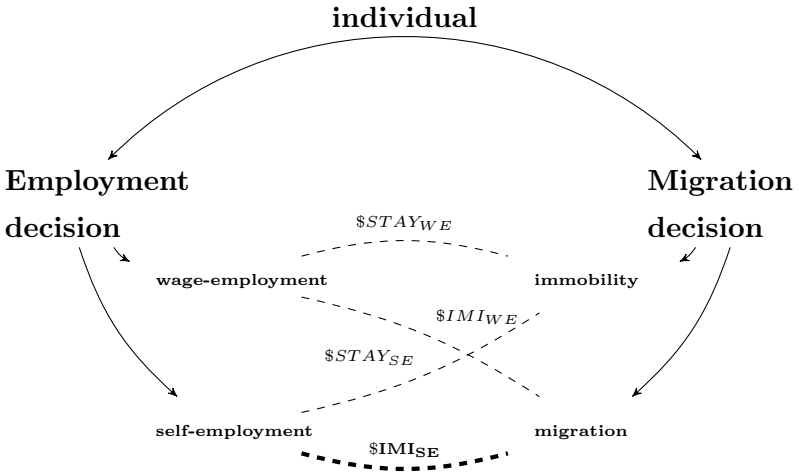
As far as undertaking self-employment is concerned, theorists in the field of the economics of entrepreneurship incline towards two complementary mechanisms of pursuing this labor market strategy. One stresses the financial benefits of choosing self-employment, be it opportunity-driven or necessity-driven. The difference being that in the first case optimization is unconstrained, and in the second case the scope of choice is limited by institutional or other impediments to employability, making the decision effectively sub-optimal. The second mechanism is related to non-pecuniary benefits of self-employment. Among these we can list self-determination or flexibility as the advantages, and increased instability and insecurity as the disadvantages.

Modern migration theorists also include non-financial benefits of emigration in the emigrants' utility function, yet their are usually limited to such features as improved living standards in the destination country, anonymity if undertaking 3-D (dangerous, dull, demeaning) jobs or access to better quality social services and social safety nets.

Despite having a common subject of study (the utility maximizing *homo oeconomicus*) and an actually observable overlap of the populations of migrants and self-employed, migration literature and self-employment literature have only moderate links to each other. Migration theories explain why people migrate, but not what form of labor market activity they choose. If they do consider the latter issue, then only in the context of the destination country, thus disregarding the initial choice of domestic or foreign economy. Theories of self-employment provide intuition as to why people choose self-employment on a given market, but do not take into consideration the fact of migrants' initial relocation.

Covering the blind spot of the two lines of conceptual developments is the key contribution of this study. The opportunity costs of immigrant self-employment considered within the conceptual framework of this study are depicted in Figure 4.1.

Figure 4.1: Comparing income gains from migration and labor market status



This chapter will propose an identification strategy and provide an answer to the question of how the optimality of the migration decision corresponds to the optimality of the form of labor market activity decision. As the subject of this research is immigrant self-employment, the possible gains to this type of labor market activity will be contrasted with immigrant wage-employment, and both wage-employment and self-employment among stayers.

From the theoretical perspective of self-employment economics, we expect that if self-employment abroad is chosen over, e.g. wage-employment at home and both options are available then the former must create an opportunity significantly more profitable (considering both financial and non-financial benefits) than the latter. Migration theories follow the same logic, despite being blind as to the specificity of labor market activity.

In an attempt to empirically test for the gains from the joint choice of self-employment and emigration, this chapter proposes the following hypotheses

Hypothesis 4.1 *Income from self-employment abroad is greater than income from self-employment at home,*

Hypothesis 4.2 *Income from self-employment abroad is greater than income from wage-employment at home,*

In order to complete the picture of self-employment as a labor market choice of immigrants, a third hypothesis will also be verified, namely that

Hypothesis 4.3 *Income from self-employment abroad is significantly different from income from wage-employment abroad.*

This general formulation of Hypothesis 4.3 results from empirical evidence provided by previous studies which showed that, depending on the context, self-employed immigrants may earn more (Borjas, 1990; Maxim, 1992) or less than their wage-employed counterparts (Portes and Zhou, 1996; Anderson, 2011). Relying solely on the pecuniary advantages of immigrant self-employment over the alternatives is mainly driven by the quantitative nature of the analysis. The concepts related to the non-monetary value of migration and self-employment will nevertheless constitute the context for possible interpretations of the results. Whether the expected income gain covers for the social, familial and cultural costs of mobility is a critical question in analyzing the migrants choice mechanism.

Despite limiting the analysis to, easily measurable, financial gains to immigrant-self-employment the verification of the hypotheses of this chapter presents several methodological difficulties. Once a person has emigrated who should we compare their income abroad to? Answering the question of whether migration is profitable requires finding a counterfactual. As in all migration studies, acknowledging the issue of non-random selection of migrants is critical. Mobility rather engages the younger, fitter, less risk-averse, and, at the same time, those with relatively poor employment or earnings prospects, from smaller towns, low-income families etc. The above enumeration of characteristics arises from the theoretical prediction that the larger the potential difference in earned incomes between the home and host countries, the more people are willing to migrate.

Migrants should thus select from those groups which are relatively poorly matched to the domestic labor market. This does not necessarily imply that potential migrants are the least able or the least educated. The contrary might just as well be the case (Borjas, 1987; Hatton and Williamson, 2004; Anacka and Okólski, 2010). At the same time potential migrants have to be employable abroad in order to generate a positive pay-off from migration. Even if this "double-selection" process was accounted for, methodologically, the intrinsic feature of migration remains an issue - that once people emigrate, they are no longer observed within the domestic population, leaving the composition of characteristics (and the correlated earnings) of the non-migrant part of the society altered by their absence.

The case of Puerto Rican emigration to the US will be considered in order to test the hypothesis that self-employed immigrants earn more than what would have earned had they not migrated or had they become wage-employed abroad. The Puerto Rico-US setting is especially well suited for this reasoning. Due to the fact that Puerto Ricans are US citizens we can neglect the positive selection of immigrants, which otherwise could have been owed to restrictive immigration policies and laws regulating employment of foreigners. Nevertheless, the spatial dimension of Puerto Rican emigration to the US remains significant, allowing to recognize Puerto Rican emigrants to the US as crossing distinct cultural, administrative and formal institutional boundaries. The case of Puerto

Rico has also been selected due to the fact that, being subject to the same data collection procedures, the Puerto Rican data perfectly matches US data, and, most importantly, similarly to US data provides information on the earnings of the self-employed. Such information is generally unavailable in the statistics of many other countries.

The next section reviews scholarly publications which addressed similar issues with the objective to determine the scope of possible empirical strategies. The following section describes the methodology used in this analysis of income gains from immigrant self-employment. Subsequently, the dataset used for the analysis is described. Next, the results, together with accompanying robustness checks, are reported. An attempt at providing external validity of the study follows in Section 4.5. The last section summarizes the outcomes and provides intuition behind the subsequent chapters of this work.

4.2 Review of research strategies

The hypothesis underlying the research on income gains from migration are strongly embedded in the neoclassical theory of migration. Based on the assumption that people migrate if they expect to experience a net income gain (Hicks, 1932; Lee, 1966; Piore, 1979; Massey, 1999; Borjas, 2000; Castles and Miller, 2009), a number of studies have made attempts at quantifying the returns to mobility. To the best of the author's knowledge all of these inquiries dealt with comparing the net income gains as measured by income from wage-employment. Nonetheless, their relevance for this study is noteworthy. Firstly, they suggest empirical strategies of dealing with the methodological challenges of estimating returns to migration. Secondly, they provide quantitative thresholds to which we will be able to compare the findings of this research.

According to Clemens (2010) the causes of international differences in wages may be rooted in one of three mechanisms. One involves a certain degree of "isolation" between markets and depends, among other things, on differences in market competitiveness, non-tradability of goods, or limitations to the transfer of technologies. The second mechanism, labeled the "growth effect" of location on human capital (Glaeser and Mare, 2001), implies that in different locations initially equally productive workers may acquire skills on the job at a different pace, e.g. due to knowledge spillovers. The third mechanism also involves externalities positively affecting one's productivity, but which augment one's level of human capital, rather then enable acquisition of new abilities (the "level effect" of location of human capital, Glaeser and Mare (2001)).

All three of these arguments can be translated into explanations for the differences in incomes of the self-employed. The first mechanism applies quite straightforward. The "growth effect" and the "level effect" of location on productivity may come from being able to mimic the best practices of other businesses, and gain or augment one's entrepreneurial performance in this manner. The self-employed are also able to gain economic advantage in one place, when compared to another, in a much simpler way, though. When considering business operation in one of a number of markets entrepreneurs may evaluate such factors

as the potential consumers' willingness to pay and the price elasticity of demand. By catering to local consumers who represent higher income levels and who are less responsive to changes in prices entrepreneurs may perform successful arbitrage. Of course their curse lies in the fact that theoretically anyone can do exactly the same, creating unfavorable competition effects.

The major methodological issue which arises when trying to compare migrants to non-migrants in terms of their earnings is that of selectivity. Migration is known to be a selective process, i.e. engaging only certain individuals from within a population. This notion poses a challenge in terms of creating relevant comparison groups among the migrants and the stayers. Whether those who leave are the "best and brightest" or the "unfit" members of the home population is a context-specific feature, depending both on the country of origin and the country of destination. Among the few empirical strategies which enable dealing with this issue are: analyses carried out in experimental settings, longitudinal studies and econometric techniques which correct the results for the extent of selectivity.

The experimental setting has been utilized in studies by Clemens (2010) and McKenzie et al. (2010). Clemens (2010) estimates wage differences among Indian workers in India and the US. Clemens took advantage of data made available by a multinational firm in the software industry which enabled workers to apply for an intra-company lottery in which they could win the opportunity to work in the American branch of the firm. The case study was chosen precisely because of the random process which determined whether people became migrants (lottery winners) or not (lottery losers). The earnings of lottery participants who won were eventually compared to the earnings of lottery participants who dod not win. Given no systematic selection into being a migrant, i.e. good balance of observable traits (education, position, age) among stayers and emigrants, and the homogeneity of the workers (tertiary educated, holding similar positions within company) enabled straightforward comparisons of mean earnings.

The results obtained by Clemens (2010) point to a 600% increase in wages when a worker moves from India to the US, despite using the same technology and producing a highly tradable good. Such results are in line with the author's conclusions from an earlier paper, namely that "gaps like these can only persist from policy barriers to labor movement. This means that those barriers create one of the largest distortions in any global market, create the largest form of wage discrimination observed in today's world" (Clemens et al., 2009, p. 56). The major drawback of Clemens' study is that it focuses on a very specific type of undermining the external validity of the findings.

McKenzie et al. (2010) have a similar strategy and analyze the income gains of emigrants from Tonga to New Zealand. The choice is also dictated by methodological considerations, i.e. the need to control for migrant selectivity. The authors take advantage of a unique dataset which holds information on all applicants for a migration lottery organized by New Zealand for the citizens of Tonga. The dataset hold information on both those who have applied for the lottery, received visas and became migrants, as well as those who

applied for the lottery, but did not receive visas and stayed in Tonga. Due to the fact that the lottery is truly random the differences in the earnings of lottery winners and lottery losers are attributable to migration its self rather than any systematic difference between those who decide to migrate and those who do not.

The authors' estimates of income gains from work abroad amount to a 263% increase in income, on average. The average weekly income of Tongans who have been successful in the migration lottery and settled permanently in New Zealand was NZD 274. The average weekly incomes of Tongans who applied for the lottery, but were unsuccessful and stayed in Tonga. The average wage of stayers is estimated at NZD 104. McKenzie et al.'s (2010) approach is methodologically very robust. Its only flaw may lay in the fact that such a setup is clearly very unique making it impossible to apply their methodology in other contexts.

Having reliable, "true" estimates based on a randomized process of selection McKenzie et al. (2010) were able to also able to compare various techniques of estimating income gains from migration. They find that single difference estimators (difference between self-reported post- and pre-migration income) overestimate the "true" income gain by 25%. Testing for this was also possible due to the specificity of the data which holds information on pre-migration incomes. A linear model estimated by OLS, which pooled migrants and non-migrants together and included a dummy for being a migrant overestimates the income gain for migrants by 31%-35%. Statistical matching techniques are also tested. Matching on multiple covariates led to a 28%-33% overestimation. Propensity Score Matching has been found to overestimate the experimental results by 20%-36%, depending on specification.

Despite potential bias, statistical and econometric techniques just as those tested by McKenzie et al. (2010) are in most cases the only way to estimate gains from migration, as experimental setups such as that in the case of Tonga or the multinational company are extremely rare. Kennan and Walker (2011) take advantage of the National Longitudinal Study of Youth and the possibility to observe exactly the same individuals over a number of years, gathering information on their earnings and place of residence. Their work differs significantly from the research reviewed heretofore in the fact that it concerns internal migration within the US. The authors consider migration as a component of a lifetime utility maximization problem, in which utility is derived from income and disutility from (costly) migration. Given the complexity of the problem on one hand, and its recursiveness on the other, the authors model it by means of dynamic programming. Though the main objective of the study was to"quantify the extent to which migration is motivated by expected income gains" (Kennan and Walker, 2011, p. 244), the model also allowed to obtain estimates of the magnitude of these gains.

The authors approximated the average gain from interstate migration in their sample at USD 1,357 per year (6.3%). When no home attachment was assumed (operationalized as a restriction on a parameter capturing a home premium) the estimated gain amounted to USD 4,424 year (20.6%), implying significant preference for staying at home. As the

authors admit, the restrictions they impose on the wage process, namely that the individual fixed effects and movements along the age-earnings profile do not affect migration decisions, are a very demanding. Due to data availability their analysis is also restricted to white, non-Hispanic, high-school graduates who did not undertake further education, leaving the question about the ability to generalize these results open.

Clemens et al. (2009) conduct the widest of existing analyses comparing the income gains from migration for emigrants from 42 countries to the US. They use a unique dataset complied by the World Bank from country censuses circa 2000 and the 5% PUMS of the 2000 US census. The wage estimates are obtained from Mincerian-type wage equations, where the dependent variable are log wages, and the set of independent variables includes standard correlates of earnings such as age, gender or residence area, and their interactions with being a US born US worker, a non-migrant (a foreign born foreign worker), an early arriver (foreign-born US worker who immigrated to the US while being under 20 years of age) or a late arriver (immigrating when older then 20). The base case model relies on a few identifying assumptions, namely that: unobserved traits of a non-migrants do not differ from those of migrants, and the relationship between these traits and wages on the home market is unrelated to the relationship of these traits and wages on the US market. Migration is also assumed to be costless.

They compare wages of observably identical migrants and non-migrants, namely 35-39 year-old, wage-employed, male workers in the urban formal sector who have 9-12 years of education in their home countries and were born there. The median purchasing power parity (PPP) adjusted gain they estimate in a base case model (without adjustments for household-level gains from having an emigrant member or personal costs of mobility) is USD 15,339. The lowest gain (37%-43%, depending on specification) is reckoned for Dominicans. Nigerian immigrants of the given characteristics are expected to gain the most, their earnings in the U.S are approximately 12 times higher than what they would have been in the home country).

As far as the issue of non-random selection of migrants is concerned Clemens et al. (2009) refute that it could be a significant source of wage-differences by comparing the unobserved components of wage-equations. Based on an analysis of few case studies they conclude that if selection was to explain the wage gaps between observably equivalent workers at home and abroad, two conditions would have to be met: 1) the workers would have to be drawn from the top few percent of the distributions of unobserved determinants of wages in their countries of origin; and 2) the unobservables which have a positive effect on wages in the home country would have to have a negative effect in the host country. It seems unlikely for these conditions to be the case. Nevertheless, the authors compute the extent of possible bias resulting from potential positive selection (on unobservable traits) of emigrants. They provide an analytical solution a potential migrant's optimization problem formulated based on Roy's (1951) model, which involves comparing foreign wages (which are known) to expected home wages (the estimated counterfactual) given the costs of migration. Within a reasonable range of parameters the estimated wage ratio overstates

the true wage ratio by 6%-18%. When quality of schooling and the imperfect transferability of formal educational attainment have been accounted for, positive selection was found to be responsible for biasing the bilateral wage ratios upward roughly by 20%-30%.

The main drawback of Clemens et al.'s (2009) research is that it provides robust estimates only for a specific category of the labor force. It also fails to account for migrants' selection into specific sectors and occupations within those sectors, what may generate a number of effects such as the domination of migrants in 3-D jobs or effective glass ceilings. And it is precisely such issues as labor market discrimination that the authors incline to in the interpretation of their results.

The review of most recent research, which provides estimates of returns to migration does not tackle the issue of returns among the self-employed. To the best of the author's knowledge no such study exists. Given the specificity of self-employment in terms of the potential risks and benefits it generates, the question of how much do people gain if they not only decide to migrate, but also optimize the form of their labor market activity seems increasingly interesting. All of the research in this area is subject to severe data limitations, an issue which will not be solved completely in this study either. For this, and other methodological reasons highlighted heretofore, all estimates of income gains to migration should be treated with caution, and with regard for the specificity of the data.

4.3 Methodology

Given the results obtained by McKenzie et al. (2010) this study will exploit propensity score matching (PSM) as a technique to generate estimates of returns to immigrant self-employment. No other empirical approach is a feasible option due to data limitations, i.e. unavailability of data collected in an experimental setting, no information on past wages, nor the possibility to apply a comparable instrumental variable to control for past income.

4.3.1 Data

The analysis uses the 5-year (2005-2010) American Community Survey (ACS) and the 5-year (2005-2010) Puerto Rican Community Survey (PRCS) data from the Integrated Public Use Microdata Series (IPUMS-USA, Ruggles et al. 2010). Using samples merged over 5-year periods implies capturing average effects over this time period[1].

All the variables from the ACS and PRCS which have been used for the purpose of estimating the relative returns to immigrant self-employment are described in Table 4.1.

The PRCS data perfectly match the ACS data as they are subject to the same data collection procedures, including the same questionnaires. This helps to avoid mistakes and imperfections of translating and merging completely distinct datasets, what would normally be the case if data sources from two different countries were used.

[1] Using 1-year samples would potentially bias the results due to the specific economic situation at that one point in time. The 2005-2010 time period allows to smooth the effects of the financial crisis, which could have been visible in the data.

Table 4.1: Description of variables used

Variable	ACS/PRCS components	Transformation
SE earnings	INCBUS00: respondent's net pre-income-tax self-employment income from a (non-incorporated) business, professional practice, or farm, for the previous calendar year; INCWAGE: respondent's total pre-tax wage and salary income for the previous year, inter alia incomes of incorporated entrepreneurs; CLASSWKRD - see the "labor market status" variable.	The two PUMS income variables have been merged and considered as self-employment income irrespective of the incorporation status of the business. Negative values were excluded from the estimations.
WE earnings	INCWAGE: respondent's total pre-tax wage and salary income for the previous year.	Used as is for wage-employed workers only.
hourly earnings	UHRSWORK: usual hours worked by the respondent per week. INCBUS00, INCWAGE - see above, CLASSWKRD - see below.	Usual hours worked per week multiplied by 52 were considered usual hours worked per year. Income (yearly) divided by the usual hours worked per year gave usual hourly income.
labor market status	CLASSWKRD: detailed class of worker, including incorporated and non-incorporated self-employment, and wage-employment. Workers with multiple sources of employment are classified according to the work relationship in which they spent the most time during the reference day or week.	Used as a dichotomous variable differentiating between the wage-employed and self-employed.
age	AGE in years.	As is.
gender	SEX: male or female.	As is.
education	EDUC: respondents' educational attainment, as measured by the highest year of school or degree completed.	Grouped into incomplete, primary, secondary, and tertiary education.
family size	FAMSIZE: counts the number of own family members residing with each individual, including the person her/himself.	Grouped in to 4 categories: individual household, family of 2 or 3, family of 4 or 5, and a family larger then 5 people.
industry	IND1990: consistent long-term classification of industries at three digit level.	Two digit classification into 14 major industry groups used.
occupation	OCC1990: consistent long-term classification of occupations at three digit level.	Two digit classification into 7 major occupational categories used.
weights	PERWT: the variable indicates how many persons in the population are represented by a given person in an IPUMS sample.	Applied to all models.

Source: The IPUMS databases, Ruggles et al. (2010).
Capitalized variable names are the names used in the IPUMS databases. For detailed distributions of the above characteristics within each of the identified groups (Table 4.2) see Table 2 in the Appendix.

Given the stated hypothesis four groups of subjects have been identified in the data: 1) self-employed Puerto Ricans residing in the US, 2) self-employed Puerto Ricans residing in Puerto Rico, 3) wage-employed Puerto Ricans residing in Puerto Rico, and 4) wage-employed Puerto Ricans residing in the US. Whenever referring to "Puerto Ricans" persons who declared their place of birth in Puerto Rico are referred to[2]. Persons considered as "residing in" the US or Puerto Rico were those captured by the ACS or the PRCS, respectively. Self-employment and wage-employment have been identified on the basis of

[2]The BPL variable in IPUMS-USA.

the self-reported class of worker. The self-employed considered in this study include both the incorporated and unincorporated entrepreneurs. The wage-employed include all other categories of workers, including those working in public institutions and the military. The sample sizes of each of the groups are reported in Table 4.2.

Table 4.2: Size of subsample used in the study, by migration and labor market status

	Frequency	Percent
Self-employed		
Puerto Ricans in Puerto Rico	4,606	57
Puerto Ricans in the U.S.	3,466	43
Total	8,072	100
Wage-employed		
Puerto Ricans in Puerto Rico	46,614	51
Puerto Ricans in the U.S.	45,075	49
Total	91,689	100

4.3.2 Propensity score estimation

Propensity score matching (PSM, Rosenbaum and Rubin (1983)) is a statistical technique which involves three steps: 1) estimating a propensity score for each individual, 2) matching observations with similar propensity scores, and 3) comparing the outcomes of matched individuals. The first and second component of the PSM procedure allow to create "statistical twins". Comparisons of people who differ in terms of one feature, but are otherwise observationally identical allows to obtain more reliable estimates of the differences in an outcome variable of interest. As matching is done on observables only, PSM makes the comparison groups structurally more alike. As indicated in the reviewed literature such an approach is one of the ways to minimize the effect of selection bias. Comparing migrants and non-migrants with similar pre-migration characteristics, rather than migrants and non-migrants as they are, increases the possibility of attributing income differences to migration.

Applying PSM results in creating a quasi-experimental setup, in which there is a "treated group" of interest, and a "control group" to which the outcomes of the treated group are compared. Following the hypotheses set for testing within this study, the groups which have been treated as experimental or control are described in Table 4.3

Table 4.3: Treated and control groups

Variable	Control group (0)	Treated group (1)
$\mathrm{SE}_{PR}SE_{US}$	self-employed Puerto Ricans	
	in Puerto Rico	self-employed
$\mathrm{WE}_{PR}SE_{US}$	wage-employed Puerto Ricans	Puerto Ricans
	in Puerto Rico	in the U.S.
$\mathrm{WE}_{US}SE_{US}$	wage-employed Puerto Ricans	
	in the US	

where SE and WE stand for self-employed and wage-employed Puerto Ricans, respectively, and PR and US stand for residence in Puerto Rico (non-migrants) and the US (migrants). An $WE_{PR}SE_{PR}$ comparison (of wage-employed and self-employed stayers) has not been considered as the fact of being a migrant in a key notion in this study. A comparison of self-employment earnings of Puerto Ricans in the US and self-employed Americans in the US also has not been performed. First, the reason is that changing one's ethnic affiliation (from Puerto Rican to American) is not a feasible choice. Second, differences in earnings of Puerto Rican immigrants in the US and of other ethnic groups will be discussed in the context of labor market discrimination in the Chapter 5.

The first methodological issue which arises when applying matching techniques is the number and type of observable characteristics which should be the basis for creating the "statistical twins"[3]. Following the premise that the more alike are the treated and control groups, the better, the matching process should be performed on all variables relevant for both the selection into the treated group, and the outcome. If one wanted to match observations on all variables as they are it could cause a computational problem due to the multidimensionality of the operation. Thus in the PSM technique Rosenbaum and Rubin (1983) suggested estimating a propensity score for each individual instead. The propensity score is the "conditional probability of assignment to a particular treatment given a vector of observed covariates" (Rosenbaum and Rubin, 1983, p. 41). They proved that adjusting for such a scalar measure is satisfactory as far as accounting for selection bias is concerned.

Technically, the estimation of propensity scores can done by running any discrete choice model. The decision which one to chose is not critical, though (Caliendo and Kopeinig, 2005). In this study a probit model is used. As a result the predicted probability of receiving treatment is obtained for every individual, given the set of one's characteristics,

[3]These do not have to be pairs of observations, as the name would indicate. Depending on the matching procedure, the "statistical twin" for an observation from the treated group could be a counterfactual statistically created on the basis of e.g. averaged characteristics of a number of observations from the control group. The procedures for creating "statistical twins" will be described in the next section.

which in this study include: age, gender, level of educational attainment, family size, one's industry of employment, and occupation (see Table 4.1 for details). For the purpose of this study propensity scores were estimated based on all possible interactions between all levels of used covariates, which included: 5 age groups, 2 genders, 2 marital statuses, 4 educational attainment levels, 4 family size groups, employment in one of 14 industries and one of 7 occupations. The interactions have been introduced in order to deal with highly inadequate matches on single covariates. This has been shown to improve the estimates obtained by means of PSM (Huber et al., 2010) and is a significant methodological improvement in comparison to how PSM was used by McKenzie et al. (2010).

An empirical strategy applying PSM, should fulfill two conditions (Caliendo and Kopeinig, 2005) - the conditional independence assumption and the common support condition. Fulfilling the common independence assumption is a major issue if one attempts to make causal inferences regarding actual the effect of treatment on the treated. This is not done in this study. The common support condition assures the comparability of the treated and control groups, what is specifically what was the reason for employing PSM in this study.

- The *Conditional Independence Assumption (CIA)* stresses that, given their characteristics, which are independent of treatment, the subjects' outcomes do not depend on whether they were treated or not. This assumption can be formally expressed as

$$(Y_0, Y_1) \perp M \mid P(X) \tag{4.1}$$

and can be translated into the requirement that the outcome (Y_0 or Y_1, e.g. income at home and abroad) is independent of the treatment (M, e.g. being a migrant), conditional on the propensity score $P(X)$. In other words the results of PSM will be valid only if we assume that if a non-migrant, who is observably identical to a migrant, moves abroad, his or her earnings will be the same as those estimated for the actual migrant. In terms of creating a quasi-experimental setup the CIA, if fulfilled, assures that while controlling for the given set of independent variables, the treatment assignment (participation in migration) is "as good as random". This is achieved under the condition that the probability of migration (being treated) is solely a function of the observable characteristics. As forceful as it is, this condition will assumedly hold in this analysis and it has been at least partly justified by the quality and quantity of data (ref: Section 4.3.1). The selection of variables used for the matching process was comprehensive. As for the fact that that the choice of covariates must be such that they are not affected by the treatment (migration), an additional clarification has to be made. Such a case is clearly true for variables such as age, gender or family size. But the effect of migration on such variables as occupation or industry might seem doubtful. Using them necessitates the assumption that emigrants work in the same occupations and in the same industries as they would at home. This is of course not necessarily true, and in many cases even highly

unlikely. Most often migrants perform jobs which are less prestigious than the ones they could perform at home. On the supply side this is possible due to the self-deprivation strategy mentioned in the introduction to this thesis. It may also be a demand-driven effect as jobs which are available for immigrants are usually from the bottom of the prestige spectrum, 3-D. Immigrants fall into those occupational categories because they have limited opportunities of obtaining legal employment at a destination and their choice is more often than not formally limited to jobs which the natives are not willing to undertake. In case of Puerto Ricans in the US this is not the case, though. As stressed at the introduction to this research Puerto Ricans have unrestricted access to the US labor market thus they face, at least formally, the same opportunity structure as "continental" US citizens. Under such circumstances it is reasonable to compare stayers and emigrants within the same occupations and industries.

- The *Common Support Condition (CSC)* requires that we actually compare treatment and control groups with the same range of propensity scores. Fulfilling this condition ensures that for each a person with a given set of characteristics both being treated or not treated are possible. The CSC can be formally expressed as

$$0 < P(M = 1|X) < 1 \tag{4.2}$$

where P is the probability measure, $M = 1$ indicates the fact the one is treated, and X is the set of covariates. The fact that it is assured that $P(M = 1|X)$ is contained in an open interval excludes cases when treatment (or not) would be perfectly predictable, leaving no reason to speculate about a possible counterfactual situation.

After imposing the CSC on the data used, the following overlap between the propensity scores of the control and experimental groups were achieved

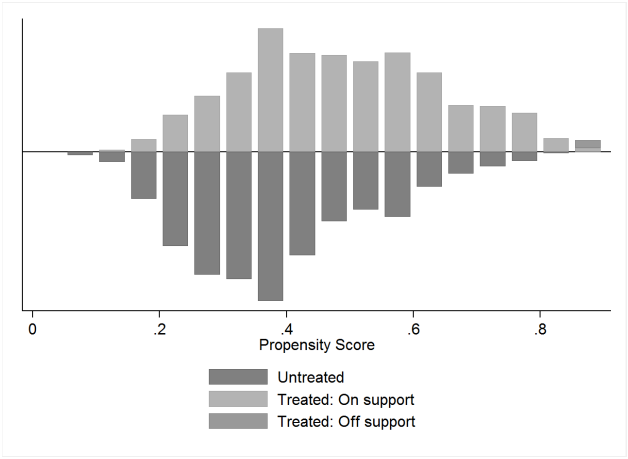
The following numbers of observations have been ignored/used due to the common support restriction:

Table 4.4 shows that after applying the common support restriction not more than three dozens of observations were lost out of the 70,000 observations used in the analysis. In all three models there were observations from the treated (migrant) group which had extremely high (incomparable with the non-migrant population) probabilities of migration (ref. Figure 4.2).

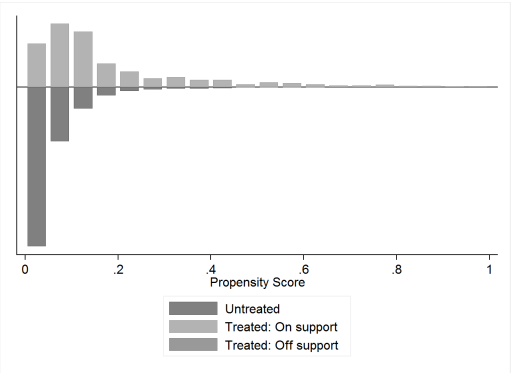
4.3.3 Matching procedures

Having estimated propensity scores for all individuals, the matching procedure is the subsequent step of implementing the PSM method. Matching is performed on the propensity scores, i.e. on the estimated probability of migrating (conditional on the set of observable characteristics) rather than on the individual characteristics themselves. This procedure

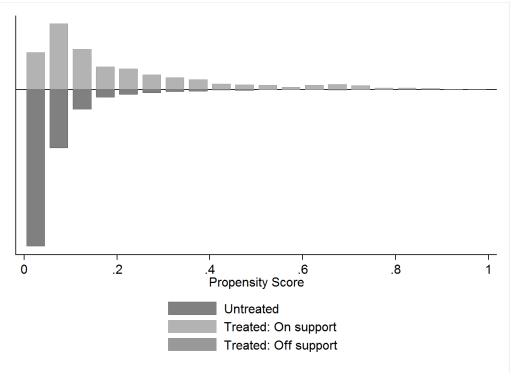
Figure 4.2: Common support



(a) SE_{PR} vs. SE_{US}



(b) WE_{PR} vs. SE_{US}



(c) WE_{US} vs. SE_{US}

Table 4.4: Common support restriction by model

Treatment	Off support	On support	Total
Self-employed in Puerto Rico and Self-employed in the US			
Control	0	3,768	3,768
Experimental	25	2,858	2,883
Total	25	6,626	6,651
Wage-employed in Puerto Rico and Self-employed in the US			
Control	0	33,177	33,177
Experimental	1	2,892	2,893
Total	1	36,069	36,070
Wage-employed in the US and Self-employed in the US			
Control	0	29,648	29,648
Experimental	5	2,888	2,893
Total	5	32,536	32,541

limits the, otherwise inevitable, multidimensionality of the matching algorithm, as mentioned in the previous section.

The aim of applying PSM in this study is to generate three counterfactual outcomes for Puerto Ricans who are actually self-employed in the US. With the objective of testing the three hypotheses posed in this chapter the following three hypothetical outcomes are

of primary interest: 1) the earnings of self-employed Puerto Rican immigrants to the US assuming that they stayed in Puerto Rico and were self-employed, 2) the earnings of self-employed Puerto Rican immigrants to the US assuming that they stayed in Puerto Rico and were wage-employed, and 3) the earnings of self-employed Puerto Rican immigrants to the US assuming that they were wage-employed in the US It is these hypothetical earnings that will be compared to the actual earnings from self-employment of Puerto Rican immigrants to the US in order to obtain robust estimates of the actual difference.

In the PSM method the difference between the outcome of the treated group and their counterfactual outcome estimated from the actual outcomes of similar individuals from the control group is called the Average Treatment Effect on the Treated (ATT), which can formally be expressed as

$$\text{ATT} = E(Y_1 \mid M = 1) - E(Y_0 \mid M = 1) \tag{4.3}$$

where Y_0 and Y_1 are average earnings of the control and experimental groups, respectively; and M is the indicator of treatment equal to 1 if one is a self-employed Puerto Rican immigrant to the US, and 0 if he or she is self-employed in Puerto Rico, wage-employed in Puerto Rico or wage-employed in the US. ATT can be translated into the difference between expected actual earnings of the treated group and their expected hypothetical earnings in the counterfactual migration-employment state.

Matching of propensity scores between the control and experimental groups may take place by means of a few different procedures including: nearest-neighbor matching, radius matching, stratification matching, local linear regression matching and kernel matching. The first three match any given observation from the treated group only with a few observations from the control group. The selection of which few observations are considered ranges from one, which is the nearest neighbor in terms of propensity score in the nearest-neighbor matching, to e.g. as many as are captured within a specified radius, in the radius matching procedure. The last two of the listed methods, kernel matching and local linear regression matching, use information on all of the observations in the control group in order to determine a statistical twin for each observation in the treated group (for a more detailed review see Becker and Ichino, 2002; Caliendo and Kopeinig, 2005). In order to test the sensitivity of the results with respect to the choice of matching procedure kernel matching, local linear regression matching and radius matching have been used. The choice was driven by the fact that kernel and local linear matching enable taking advantage of all the information in the dataset. These two methods do not leave any doubt as to which observations get matched and which do not, what is the case in the other procedures. Radius matching was chosen for sensitivity analysis, precisely because it involves a great amount of arbitrariness in determining which observations are considered as reasonable matches, in what it greatly differs from kernel and local linear matching. The matching procedures were bootstrapped with 1000 replications each in order to correct for the fact

that the standard errors of the estimated outcomes do not take into consideration the fact that that propensity scores were calculated beforehand[4].

In kernel matching a counterfactual outcome for every person in the experimental group is constructed as a weighted average of the outcomes of persons from the control group. Greatest weight is given to subjects with closest propensity scores. The weighting is done by means of a kernel function. For the purpose of this study a bi-weight kernel has been chosen as it best rewards better matches. The bandwidth of the kernel was left at the default value. Counterfactual outcomes ($E(Y_0 \mid M = 1)$, see Equation 4.3) estimated by means of kernel matching can be obtained by running a weighted least squares regression of Y_0 on an intercept, where the weights are the kernel weights, $K(\frac{P(X_i)-P(X_j)}{b})$, and where K is the kernel function, b is the bandwidth parameter, and P_i and P_j are the propensity scores of individuals from the experimental and control groups, respectively. The estimate of the counterfactual mean is the estimated intercept from such a regression.

Estimating the counterfactual obtained from local linear regression matching differs from that from kernel matching in that, on top of the intercept, one adds a linear term in P to the regression equation. Smith and Todd (2005) show that such an adjustment helps the estimator to be robust to cases in which the observation from the control group are distributed asymmetrically around the experimental observations

Radius matching matches observations from the experimental group with those from the control group which are nearest neighbors and which satisfy that condition that they fall within a given distance between P_i and P_j. This has been a significant improvement to simple Nearest Neighbor Matching which found matches regardless of how the two observations were actually different. Radius Matching avoids bad matches. In this sense it is also of imposing the Common Support Condition (Smith and Todd, 2005). Nevertheless, Smith and Todd (2005) mention that it may be difficult to deduce what interval should be chosen as feasible for making matches. The results presented in this chapter employ a caliper equal to 1 standard deviation (s.d.) of the propensity scores estimated for a the given sample. The results of sensitivity analysis considering 0.5 s.d. and 2 s.d. calipers are presented in the Appendix.

After performing the matching procedures it was tested whether the average levels of input variables become more similar between the experimental and control group in comparison to the unmatched samples, i.e. whether the control and treated groups are structurally alike. The results are presented in Table 4.5.

4.3.4 Balancing properties

The results of tests for the equality of input characteristics (propensity score estimation covariates) slightly vary across comparison groups (see Table 4.5). Self-employed Puerto

[4]The *psmatch2* program was used for matching on the propensity scores and comparing the outcome variables. For each of the three models computations involving every of the three matching procedures lasted approximately 6 days each on a 4GB RAM computer with a dual core processor and Stata 12 software.

Ricans in Puerto Rico and self-employed Puerto Ricans in the US are very well matched. On neither level of the any of the covariates does the difference in means exceed 5%, the level considered acceptable in the literature (Caliendo and Kopeinig, 2005). This is in line with the expectation that comparing people with the same labor market status is potentially less prone to selectivity bias. When wage-employed stayers are compared to self-employed migrants the greatest number of imperfect matches is visible (values in bold in Table 4.5). Five covariates maintain bias exceeding an acceptable 5% level. These variables are age in the range 40-50 years, marital status, work in the manufacturing and retail trade industries, and work as a technician, salesman or administrative worker. The t-test forces us to reject the hypothesis of the equality of their means when comparing the treated group to the matched control group. From Table 4.5 it follows that as far as the 40-50 age group and married persons are concerned, they are overrepresented among the migrants. For the two industries the opposite is true - workers in manufacturing and retail are overrepresented among the stayers. Workers in the technical, sales, and administrative occupations are also overrepresented among migrants. In the last analysis, where self-employed Puerto Rican migrants residing in the US are compared to wage-employed migrants, only one variable remained slightly biased - employment in the manufacturing industry was slightly more common among the wage-employed. Thus it is visible that the greatest imperfections in the balancing of covariates occur when we consider groups which differ in both their of labor market activity and migration status. .

Despite some bias still remaining after matching, especially between self-employed migrants and wage-employed stayers, it is worth noting that all but one of the differences in means which were statistically significant were smaller than 2 percentage points. The proportion of married self-employed migrants after exceeds after matching that of married wage-employed stayers by 3.2 percentage points. The inability to obtain perfect statistical twins in the second analysis arises from the, rather unsurprising, differences in market structures of Puerto Rico and the US, the selectivity of entrepreneurs with respect to those industries in which small- and micro-entrepreneurship is feasible, and the selectivity of migration its self. More detailed diagnostics of the balancing properties showed that in comparison to the unmatched samples the reduction in bias on all of the significantly different covariates was very large, around 70%. The overall mean absolute standardized bias in the SE_{PR}/SE_{US} groups was reduced to 2.17%; in the WE_{PR}/SE_{US} groups to 2.82%; and in the WE_{US}/SE_{US} groups to 2.63%. These result leave us confident about the feasibility of further comparisons of the outcome variables across these groups.

Table 4.5: Balancing properties (after kernel matching)

	Self-employed in the US vs. Self-employed in Puerto Rico					Self-employed in in the US vs. Wage-employed in Puerto Rico					Self-employed in the US vs. Wage-employed in the US				
	Mean			t-test		Mean			t-test		Mean			t-test	
Variable	Treated	Control	%bias	t	$p > t$	Treated	Control	%bias	t	$p > t$	Treated	Control	%bias	t	$p > t$
age 20-30	0.22	0.22	-0.1	-0.05	0.96	0.22	0.22	-0.2	-0.08	0.94	0.22	0.23	-2.1	-0.78	0.44
age 30-40	0.28	0.28	0.5	0.17	0.87	0.28	0.28	1.2	0.44	0.66	0.28	0.29	-1.8	-0.67	0.50
age 40-50	0.26	0.26	1.7	0.63	0.53	0.26	0.24	5.4	1.97	**0.05**	0.26	0.25	3.6	1.37	0.17
age 50-60	0.14	0.15	-2.4	-0.91	0.36	0.14	0.13	3.3	1.13	0.26	0.14	0.13	3.3	1.23	0.22
gender-female	0.31	0.30	2.5	0.9	0.37	0.31	0.33	-2.4	-0.96	0.34	0.31	0.34	-4.8	-1.87	0.06
education-primary	0.08	0.09	-2.6	-1.12	0.26	0.08	0.08	0.5	0.17	0.86	0.08	0.08	2	0.69	0.49
education-secondary	0.38	0.40	-2.8	-1.06	0.29	0.38	0.37	1.7	0.66	0.51	0.38	0.38	-0.1	-0.02	0.98
education-tertiary	0.52	0.50	3.9	1.48	0.14	0.52	0.54	-2.8	-1.07	0.28	0.52	0.53	-1.4	-0.51	0.61
married-yes	0.65	0.63	3.1	1.2	0.23	0.65	0.61	8.3	3.21	**0.00**	0.65	0.63	4.8	1.83	0.07
family 2-3	0.44	0.44	0.5	0.19	0.85	0.44	0.44	1	0.36	0.72	0.44	0.44	-0.3	-0.11	0.92
family 4-5	0.32	0.31	2.6	0.96	0.34	0.32	0.32	0.3	0.1	0.92	0.32	0.31	1.5	0.57	0.57
family 5 <	0.05	0.05	0.9	0.35	0.73	0.05	0.06	-1.4	-0.51	0.61	0.05	0.06	-2.1	-0.77	0.44
construction	0.16	0.17	-2.2	-0.88	0.38	0.16	0.15	4	1.37	0.17	0.16	0.16	2.1	0.66	0.51
manufacturing	0.03	0.04	-2.9	-1.03	0.30	0.03	0.05	-8.1	-4.03	**0.00**	0.03	0.05	-6.1	-2.85	**0.00**
transport/communications	0.05	0.06	-2.1	-0.75	0.46	0.05	0.06	-2.6	-1.01	0.31	0.05	0.06	-2.6	-1.08	0.28
wholesale	0.03	0.03	1.8	0.6	0.55	0.03	0.03	2.1	0.84	0.40	0.03	0.03	1.7	0.62	0.54
retail	0.15	0.17	-3.5	-1.34	0.18	0.15	0.17	-6	-2.33	**0.02**	0.15	0.15	1.5	0.52	0.61
finance/insurance	0.05	0.04	5.3	1.84	0.07	0.05	0.05	0.4	0.15	0.88	0.05	0.05	0.6	0.25	0.81
business/repair services	0.12	0.12	1	0.38	0.71	0.12	0.11	4.4	1.5	0.13	0.12	0.11	5.3	1.78	0.08
personal services	0.07	0.07	-1.4	-0.55	0.58	0.07	0.07	2.1	0.7	0.49	0.07	0.06	3	1	0.32
entertainment	0.01	0.01	0.6	0.2	0.84	0.01	0.01	-0.7	-0.22	0.82	0.01	0.02	-3.1	-1.07	0.29
professional services	0.26	0.24	4.1	1.45	0.15	0.26	0.25	1.9	0.76	0.45	0.26	0.28	-4.5	-1.83	0.07
technician, salesman, administrative worker	0.21	0.21	1	0.36	0.72	0.21	0.23	-5	-2	**0.05**	0.21	0.22	-2.1	-0.83	0.41
service worker	0.16	0.16	-1.5	-0.56	0.58	0.16	0.15	1.9	0.73	0.46	0.16	0.15	1.4	0.55	0.58
farmer etc.	0.05	0.05	0.8	0.36	0.72	0.05	0.04	2.7	0.86	0.39	0.05	0.04	5.2	1.63	0.10
precision or repair worker, craftsman	0.14	0.15	-2.1	-0.84	0.40	0.14	0.14	1.2	0.41	0.68	0.14	0.14	0.1	0.02	0.98
operator, fabricator, laborer	0.11	0.13	-4.8	-1.89	0.06	0.11	0.13	-4.5	-1.77	0.08	0.11	0.13	-4	-1.57	0.12

4.4 Results: Income gains from immigrant self-employment

The considered outcome variables were hourly and yearly incomes from self-employment or wage-employment of migrants or stayers, depending on the hypothesis being tested. This section presents the results of comparisons of the average levels of actual earnings of self-employed migrants with their counterfactual earnings had they been in the control group, i.e. the Average Treatment Effect on the Treated (ATT). The ATT will be thus considered as an estimate of the returns to immigrant self-employment. Following the the conventional way of presenting results in this study, the first part of Table 4.6 compares gains from Puerto Rican immigrant self-employment in the US in case when self-employment in Puerto Rico is considered an alternative, the second part of Table 4.6 treats wage-employment in Puerto Rico as an employment alternative and the last part of Table 4.6 compares income gains of Puerto Rican self-employed to Puerto Rican wage-employed in the US.

Table 4.6: Income gains to migration: Average Treatment Effects on the Treated (ATT)

	Before matching		After matching							
Variable	Treated	Controls	Treated	Controls	Difference	PSM S.E.	T-stat	Bootstrap S.E.	z	$P > \|z\|$
Self-employed in Puerto Rico vs Self-employed in US										
Kernel Matching										
hourly income	18.81	11.42	18.82	14.12	4.70	0.68	6.92	0.77	6.26	0.00
yearly income	38,953	18,911	38,939	23,602	15,337	1,254	12.23	1,271	12.27	0.00
Local Linear Matching										
hourly income	18.81	11.42	18.82	13.99	4.83	0.69	7.05			
yearly income	38,953	18,911	38,939	23,336	15,603	1,261	12,38			
Radius Matching										
hourly income	18.81	11.42	18.82	11.77	7.06	0.63	11.18			
yearly income	38,953	18,911	38,939	19,584	19,355	1,206	16.05			
Wage-employed in Puerto Rico vs Self-employed in US										
Kernel Matching										
hourly income	18.77	10.94	18.78	12.04	6.74	0.53	12.64	0.53	12.66	0.00
yearly income	38,884	20,950	38,887	23,789	15,099	1112	13.58	1,128	13.38	0.00
Local Linear Matching										
hourly income	18.77	10.94	18.78	12.09	6.69	0.54	12.50			
yearly income	38,884	20,950	38,888	23,945	14,943	1114	13.42			
Radius Matching										
hourly income	18.77	10.94	18.78	11.77	7.00	0.53	13.2			
yearly income	38,884	20,950	38,888	23,147	15,741	1,109	14.20			
Wage-employed in U.S. vs Self-employed in US										
Kernel Matching										
hourly income	18.77	15.67	18.78	17.25	1.53	0.59	2.58	0.61	2.51	0.00
yearly income	38,884	30,808	38,912	34,838	4,074	1,137	3.58	1,169	3.49	0.00
Local Linear Matching										
hourly income	18.77	15.67	18.78	17.58	1.20	0.60	2			
yearly income	38,884	30,808	38,912	35,531	3,381	1,142	2.96			
Radius Matching										
hourly income	18.77	15.67	18.78	16.82	1.96	0.58	3.38			
yearly income	38,884	30,808	38,912	33,864	5,048	1,129	4.47			

The caliper size assumed for Radius Matching was 1 s.d. of the estimated propensity scores for a given sample.

Results with other calipers, 0.5. s.d. and 2 s.d., are presented in the Appendix.

The alternative specifications do not alter the results significantly.

The results show a consistently positive and significant Average Treatment on the Treated, implying positive gains to immigrant self-employment in comparison to all three of the considered counterfactuals. Table 4.6 is indicative of slight differences of magnitude between estimates obtained by means of kernel matching or local linear regression matching and radius matching, though. Testing for average bias reduction after all three matching methods lead to conclude that average bias is lowest after kernel matching. Thus estimators of income differences obtained using this method will be subject to further interpretations.

The comparison of earnings of the self-employed in Puerto Rico and in the US shows significant gains to migration. On average, a Puerto Rican entrepreneur earns nearly USD 5 per hour more in the US than he or she would earn in Puerto Rico. This is a significant difference given that it amounts to 70% of the US federal nationwide minimum wage level of USD 7.25 per hour, and it is more then the Puerto Rican minimum wage of USD 4.10. The difference in yearly income from self-employment between stayers (the control group) and emigrants (the treated group) amounted to approximately USD 15,000. After bootstrapping the ATT estimates our initial results remain unchanged. This is more then USD 1,000 per month. In percentage terms the USD 15,000 gain amounts to 65% of the estimated earnings from self-employment in Puerto Rico. Recall that the income gain for Puerto Rican migrants to the US obtained by Clemens et al. (2009) was of comparable magnitude, 50%. From the obtained results it follows that just by migrating people are able to significantly improve their financial situation.

Having considered the simplest case when the income gain to immigrant self-employment results purely from migration and not changing one's labor market status, an exercise combining both effects was considered. The second part of Table 4.6 presents a comparison of earnings of self-employed Puerto Rican immigrants in the US and the earnings of wage-employed non-migrants remaining in Puerto Rico. When looking at the ATT estimate when the profitability of self-employment in the US is compared to the profitability of wage-employment in Puerto Rico, the income gains from the former strategy are also evident on the level of both hourly and yearly earnings. As far as yearly earnings are concerned the gain is similar to that experienced by Puerto Rican who were assumed to be self-employed prior to migration, approximately USD 15,000. When it comes to hourly earnings the gain is even larger. Self-employed Puerto Rican emigrants to the US earn approximately USD 7 more than if what they would earn if they were wage-employed at home.

When the gains from transitioning from wage-employment are compared to the gains from transitioning from self-employment we can come to a conclusion that the supply of labor of these two potential groups of emigrants differs. Those who migrated and changed their labor market status gain more per hour, but in yearly terms their earnings are nearly identical to those of people who migrated but did not change their labor market status. If it is that the self-employed devote more time to economic activity than do wage-employed, than one possible narrative behind this result could be that the those who were wage

employed in Puerto Rico did not change their work schedule. Earning relatively more but working relatively fewer hours could lead to identical gains as those of people who earn less, but work more.

From the comparison of self-employed migrants to wage-employed or self-employed stayers we can deduce that emigration to the US and taking up self-employment is an income-maximizing choice for Puerto Ricans. If they weigh their gains to immigrant self-employment against the option of being self-employed in Puerto Rico they may gain per year as much (approximately USD 15,000) as they would earn in seven and a half months back at home (USD 23,602 per year, approximately USD 2,000 per month). The cost for achieving this is family separation and potentially also poorer living conditions, unaccounted for in this study. If gains to immigrant self-employment are weighed against the option of being wage-employed in Puerto Rico the yearly gain is just as high. If one wanted and could spend more time on the job, though, the benefit could be even higher in this case.

When we compare gains to self-employment and gains to wage-employment among Puerto Ricans in the US, the findings are less optimistic, yet the gain is still positive (last part of Table 4.6). It appears that self-employed Puerto Rican immigrants earn significantly more than their US-based wage-employed statistical twins. In terms of hourly earnings the difference is approximately USD 1.5, while in yearly terms it amounts to circa USD 4,000. These final outcomes reaffirm the income-maximizing objective of self-employed Puerto Rican migrants to the US.

4.5 External validity of the study

In order to provide a benchmark for the obtained results of the case study of Puerto Rican migration to the US a similar analysis has been carried out on a different sample of migrants. Due to the preferential position of Puerto Ricans on the US labor market, i.e. lack of formal restriction to labor market participation, the closest to migration from Puerto Rico to the US was internal migration across US states.

The dataset used was the same one as that in the above Puerto Rican study, yet only its US part, i.e. the 5-year 2010 American Community Survey (Ruggles et al., 2010). The variable used to differentiate movers from stayers (MIGRATE1) reports whether an individual had changed his or her residence since a reference point one year ago. The state in which the person was enumerated (STATEFIP) was compared to the state of residence one year ago (MIGPLAC). If the distance between the capital cities of those states was larger than 1000 miles (the distance from San Juan, Puerto Rico to Miami, Florida), the person was considered a mover. If one declared he or she did not change the state of residence within the past year, they were considered a stayer. The basic descriptive statistics for both groups are presented in the Appendix.

As the self-employed are of main interest to this study, only self-employed movers and self-employed stayers were compared. The propensity score related to being a mover

was computed for each individual in the final sample. The model used to estimate the propensity score was identical to that applied in the Puerto Rican study. Due to the fact that that migration was restricted to mobility exceeding the distance of 1000 miles and to persons who were self-employed, the proportion of movers was very small, 0.45%. Because of this and the fact that the distributions of propensity scores in both samples were similar, one-to-one matching was applied rather than techniques which weighed every observation in the control population. One-to-one matching is a version of nearest neighbor (NN) matching. The method matches an individual from the treated group to an individual from the comparison group that is closest in terms of the propensity score. Only one nearest neighbor, the default, has been considered for the base case comparison[5]. The procedure can be carried out in a way that already matched individuals from the comparison group can be used more then once (matching with replacement) or not (matching without replacement). Replacement was allowed, what lead to increased quality of matching and smaller bias across the samples of movers and stayers. The test of the balancing properties of the predictor variables are presented in Table 4.7. In all cases remaining bias is negligible.

After establishing a good control group for internal migrant entrepreneurs, the earnings of the two groups were compared. Analogously to the Puerto Rican study, both hourly and yearly income was considered. The estimated differences in returns to self-employment for people who are internally mobile and those who are not are presented in Table 4.8. In order to test for the robustness of this result one additional attempt has been made to understand the gains to self-employment in case of internal migration in the US. The lower section of Table 4.8 also presents the results obtained by the same method as the one applied above, but where the outcomes to be compared were normalized incomes, rather than absolute incomes from self-employment. Every individual's income from self-employment has been divided by the mean earnings from self-employment in the state if his or her residence.

Table 4.8 indicates that both in the case of hourly and yearly earnings internal migrants (the treated group) earn significantly less, than their immobile statistical twins. In case of hourly earnings the difference is USD 3.5, while in terms of yearly income the difference amounts to over USD 10,000. This finding contradicts the intuition that people move to improve their material status.

[5]A matching procedure taking into account 5 nearest neighbors has also been applied as a form of sensitivity analysis. The results were similar in terms of sign and significance, but smaller in magnitude. The balancing properties were satisfied.

Table 4.7: Balancing properties, internal migration (after one-to-one matching)

		Mean				t-test	
Variable	Sample	Treated	Control	%bias	bias	t	p>t
age 20-30	Unmatched	0.27	0.17	22.5		13.84	0.00
	Matched	0.27	0.27	-0.3	98.7	-0.11	0.91
age 30-40	Unmatched	0.23	0.28	-13.2		-7.20	0.00
	Matched	0.23	0.23	-0.2	98.4	-0.09	0.93
age 40-50	Unmatched	0.20	0.30	-24.5		-12.99	0.00
	Matched	0.20	0.20	0.1	99.4	0.06	0.95
age 50-60	Unmatched	0.11	0.17	-18.1		-9.41	0.00
	Matched	0.11	0.11	0.2	99.0	0.08	0.94
gender-female	Unmatched	0.40	0.36	8.3		4.72	0.00
	Matched	0.40	0.40	0.1	98.4	0.05	0.96
education-primary	Unmatched	0.02	0.03	-5.6		-2.93	0.00
	Matched	0.02	0.02	0.2	96.4	0.09	0.93
education-secondary	Unmatched	0.30	0.39	-20.7		-11.36	0.00
	Matched	0.30	0.30	-0.2	99.0	-0.08	0.94
education-tertiary	Unmatched	0.68	0.57	22.4		12.31	0.00
	Matched	0.68	0.68	-0.1	99.7	-0.03	0.98
married-yes	Unmatched	0.54	0.72	-37.5		-22.45	0.00
	Matched	0.54	0.55	-0.3	99.1	-0.13	0.90
family 2-3	Unmatched	0.45	0.51	-12.9		-7.31	0.00
	Matched	0.45	0.45	-0.1	99.0	-0.05	0.96
family 4-5	Unmatched	0.17	0.26	-22.7		-11.96	0.00
	Matched	0.17	0.17	-0.1	99.7	-0.03	0.97
mining	Unmatched	0.04	0.05	-6.0		-3.17	0.00
	Matched	0.04	0.04	0.3	94.8	0.13	0.89
construction	Unmatched	0.00	0.00	1.2		0.72	0.47
	Matched	0.00	0.00	0.7	44.7	0.26	0.80
manufacturing	Unmatched	0.15	0.18	-7.6		-4.18	0.00
	Matched	0.15	0.15	0.1	98.9	0.03	0.97
transport/communications	Unmatched	0.04	0.04	-2.1		-1.17	0.24
	Matched	0.04	0.04	0.2	92.7	0.06	0.95
wholesale	Unmatched	0.04	0.05	-0.6		-0.31	0.76
	Matched	0.04	0.04	0.0	100.0	-0.00	1.00
retail	Unmatched	0.02	0.03	-2.0		-1.09	0.28
	Matched	0.02	0.024	0.2	90.0	0.08	0.94
finance/insurance	Unmatched	0.10	0.11	-3.4		-1.90	0.06
	Matched	0.10	0.10	-0.2	94.2	-0.08	0.94
business/repair services	Unmatched	0.09	0.08	1.7		0.96	0.33
	Matched	0.09	0.09	-0.3	80.2	-0.13	0.90
personal services	Unmatched	0.13	0.11	5.5		3.20	0.00
	Matched	0.13	0.13	-0.1	98.2	-0.04	0.97
entertainment	Unmatched	0.08	0.09	-0.7		-0.41	0.68
	Matched	0.08	0.08	-0.1	84.6	-0.05	0.96
professional services	Unmatched	0.02	0.01	8.7		5.93	0.00
	Matched	0.02	0.02	0.5	94.6	0.17	0.87
public administration	Unmatched	0.27	0.21	13.6		8.02	0.00
	Matched	0.27	0.27	-0.1	99.5	-0.03	0.98
technician, salesman, administrative worker	Unmatched	0.21	0.22	-0.5		-0.30	0.76
	Matched	0.21	0.21	0.3	43.1	0.12	0.90
service worker	Unmatched	0.13	0.14	-2.9		-1.64	0.10
	Matched	0.13	0.13	-0.2	93.7	-0.07	0.94
farmer etc.	Unmatched	0.04	0.09	-18.5		-9.09	0.00
	Matched	0.04	0.04	0.5	97.2	0.25	0.80
precision or repair worker, craftsman	Unmatched	0.12	0.15	-9.1		-4.93	0.00
	Matched	0.12	0.12	-0.1	99.0	-0.04	0.97
operator, fabricator, laborer	Unmatched	0.07	0.08	-4.0		-2.20	0.03
	Matched	0.07	0.07	-0.2	94.1	-0.10	0.92

Table 4.8: Average treatment effect on the treated, internal migration

	Before matching		After matching				
Variable	Treated	Controls	Treated	Controls	Difference	S.E.	T-stat
Income as is							
hourly income	21.29	25.13	21.29	24.75	-3.45	1.53	-2.25
yearly income	40,269	51,115	40,269	50,725	-10,456	3,163	-3.31
Normalized income							
hourly income	0.85	1.03	0.85	1.00	-0.15	0.06	-2.50
yearly income	0.91	1.16	0.91	1.13	-0.22	0.07	-3.13

The above results confirm that internally mobile entrepreneurs earn significantly less than stayers who are self-employed. Both in case of hourly and yearly earnings movers

earn below the average, while stayers earn at (hourly income) or above (yearly income) the average. The earnings of movers are 85% of those of the stayers in hourly terms, and 80% in case of yearly earnings. This contradicts the findings of Kennan and Walker (2011), who found significant gains to internal migration in the US.

Two reasons may underlie this conclusion. First, the data limits the tome horizon in which mobility is observed to one year. This implies that the self-employment observed among movers is very new. It is easy to imagine that in a new start-up the owners earnings are relatively small to what all those who did not move in that past year, thus on average have well established businesses, earn. Second, in case of internal migration we cannot exclude motives other than financial gains.

The US Current Population Survey[6] provides information on the motives of internal mobility. In 2012, among movers of all ages, 23.23% of the 3 million long-distance intra-U.S. mobility (exceeding 500 miles) was due to family-related reasons. 52.04% (1.5 million persons) moved for employment-related reasons, including new job or job transfer, to look for work or lost job, to be closer to work/easier commute or becoming retired. 22.30% of movements were predominantly for housing-related reasons, 2.42% was mobility for other reasons. Thus, it might be the case that for a significant proportion of the mobile population (circa 50%) self-employment is a by-product of their move, rather than its objective. In this sense it may be considered as a necessity rather than an opportunity. Such businesses are likely to be less profitable. Neither of these reasonings can be tested on the data used in this study. Nevertheless, they constitute interesting avenues for future research and possibly could be considered on the basis of the Current Population Survey.

4.6 Summary and conclusions

Stemming from migration theory is the presumption that people migrate due to relative wage differences between the sending and receiving markets. This postulate is predominant in neoclassical theories, but modern concepts of migration do not refrain from acknowledging it either. Among the few studies which make attempts at evaluating return to migration neither, to the best of the author's knowledge, deals with gains to immigrant self-employment, as opposed to wage-employment. Due to the multilaterally beneficial role of immigrant self-employment the economic rationality behind choosing this labor market strategy calls for scientific attention.

This chapter aimed at testing whether immigrant self-employment is an income-maximizing strategy when compared to either of the alternatives: wage-employment at home, self-employment at home, or wage-employment in the host country. Given the data limitations on one hand, and the ability to eliminate policy driven selection bias on the other hand, migration from Puerto Rico to the United States has been the subject of the study.

[6] www.census.gov/hhes/migration/data/cps/cps2012.html

The estimates of income gains revealed the following relations of incomes from immigrant self-employment and three possible alternatives, i.e. self-employment at home, wage-employment at home, and wage-employment abroad

in hourly terms:

$$SE_{US} > SE_{PR} \text{ by USD } 4.70 \tag{4.4a}$$

$$SE_{US} > WE_{PR} \text{ by USD } 6.74 \tag{4.4b}$$

$$SE_{US} > WE_{US} \text{ by USD } 1.53 \tag{4.4c}$$

and in yearly terms:

$$SE_{US} > SE_{PR} \text{ by USD } 15,337 \tag{4.4d}$$

$$SE_{US} > WE_{PR} \text{ by USD } 15,099 \tag{4.4e}$$

$$SE_{US} > WE_{US} \text{ by USD } 4,074 \tag{4.4f}$$

The financial gains to immigrant self-employment, researched in the context of Puerto Rican immigration to the US, are unambiguously positive. This conclusion holds for all three cases considered - when the alternative is wage-employment at home, self-employment at home or wage-employment at the destination. The presented extension to the Puerto Rican study revealed that such gains need not to be taken as granted, though. Extending the analysis to mobility of comparable distance across US states showed that self-employed migrants may earn less than than self-employed stayers, i.e. that migration can be related to a decrease in earnings. The data limitations of this study could have lead to a downward bias of these results, though. Further research in other settings still remains to be done.

On top of the financial gains, as noted in the literature review to this thesis, non-financial costs and benefits are imperative in the assessment of the drivers of migration. This is to say that for migration to be a feasible option it must not only promise to enable higher earnings per se, but also cover for the possible socio-psychological costs of mobility. The differences in earnings would thus have to be higher if one assigns negative value to e.g. separation with family or to the (possibly demeaning) job one is to undertake abroad. Of course the destination country may also offer conditions which are beneficial for the migrant and may decrease the difference in real earning necessary for migration to occur. Such conditions include, but are not limited to, better living standards, health care or child care, social benefits. The destination country may also offer anonymity. This becomes valuable if one is to perform a 3-D job, which one would never undertake at home in fear of humiliation. Matching on variables which could be considered proxies of such features and then comparing obtained estimated of income differences could be a research strategy for yet another extension to this study.

Given the estimated pecuniary gains and the possible non-pecuniary gains from migration the obtained results beg the question of whether they cover the socio-economic costs of mobility. Even when considering only the financial cost related to, among other things,

the price for the journey a migrant would have to consider spending approximately USD 500 on a 2.5hr round trip flight from San Juan, Puerto Rico to Miami, Florida[7]. This decreases the yearly income gain for migrants by merely 3%, to USD 14,500, an amount of money one most probably could and would spend on occasional visits home without loosing the financial motivation of being a self-employed migrant.

The finding that immigrant self-employment is a unambiguously profitable option when compared to wage-employment at home or abroad or self-employment at home is surely not the only reason why individuals choose that labor market strategy. Actually a number of other push and pull factors may also play a decisive role in the process. If emigrants consider changing labor market status while being abroad, one of the underlying aspects of the decision-making could also be the possibility to avoid discriminatory treatment on the labor market. An inferior position in comparison to one's group of reference might be an incentive strong enough to push people out of wage-employment into self-employment, where the returns depend on one's productivity to a much greater extent. On the contrary, when choosing immigrant self-employment as an alternative to domestic self-employment, the decision-making process may account for factors which make migration a superior option to staying, thus pulling entrepreneurs abroad. Among such factors we can identify ethnic enclaves as an institution which possibly contributes to the difference in returns to self-employment at the origin and at the destination countries. Absent at home but present where one's ethnic group constitutes a minority population, it is said that ethnic enclaves may serve as business incubators in which immigrant businesses can achieve returns unattainable otherwise. Both labor market discrimination, as a potential push factor, and ethnic competition and complementarity in ethnic enclaves, as a potential pull factor, will be the subject of study in the following chapters.

[7]Results of search for the cheapest return flight in a month's time on expedia.com.

Chapter 5

Discrimination

5.1 Introduction

Chapter 4 showed that immigrant self-employment can bring relatively high returns, even when compared to wage-employment abroad. This feature implies that undertaking entrepreneurship may be a very good way for migrants to deal with labor market imperfections, such as wage discrimination. Discrimination may be very distressful for those who are affected. On one hand it is quite straightforwardly related to a penalty in terms of earnings. On the other hand, unfair treatment also causes psychological tension. In case of immigrants lack of agency and empowerment, especially in the lower segments of the labor market may conceal these detrimental effects. As it will be argued, self-employment may be a way for immigrants to deal with the issue of discrimination.

One of the types of discrimination which underlies the choice between wage-employment and self-employment, is employer discrimination. Employers may have a distaste for employing immigrants, i.e. they may have an exogenous taste for discriminating against them. Employer discrimination can be manifested in two ways - when people with similar productivity characteristics are payed differently, i.e. *wage discrimination*, and/or when they are unable to obtain similar positions, i.e. *occupational discrimination*.

Differences between groups in observed labor market outcomes, such as earnings, may be due to two different factors, though. On one hand, it may be the fact they are differently endowed in productivity-related features, e.g. have different levels of school attainment or labor market experience. On the other hand, it may be because of how differently their endowments are evaluated by employers. The first source of differences in earnings would not be considered discriminatory or unfair. If it happens that due to the selectivity of migration poorly educated and poorly skilled migrants arrive at the destination, then their lower earnings may adequately reflect their productivity. If these workers nevertheless earn less than natives with equally poor skills and abilities, then that is what we would recognize as discrimination. Thus what is critical in researching labor market discrimination is distinguishing between the two factors responsible for wage differences.

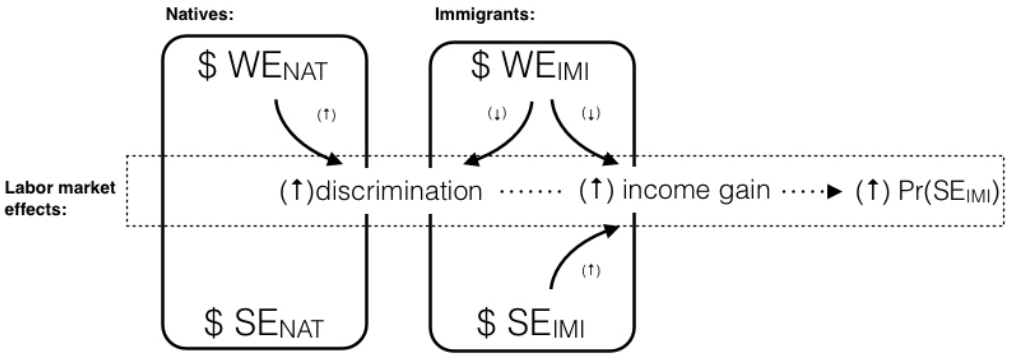
From the mechanisms described above it follows that immigrants either may have limited access to higher payed positions, or they are paid less. Whichever is in place it

makes wage-employment relatively less attractive compared to self-employment. For some individuals this may be an incentive strong enough to undertake the latter. The theory of employer discrimination suggests that the push-pull factors towards self-employment are created by the following inequality

$$w_M/w_N < \pi_M/\pi_N$$

This indicates that as long as their relative earnings from self-employment exceed their relative earnings from wage-employment, migrants might concentrate in the former. Crowding into self-employment is hypothesized to continue until the profits are pushed down to a point where the relative earnings of immigrants in both types of labor market activity are equal. The interdependencies between immigrant and native labor market activity and immigrants' intensity of self-employment is conceptually depicted in Figure 5.1.

Figure 5.1: Immigrant discrimination and self-employment intensity

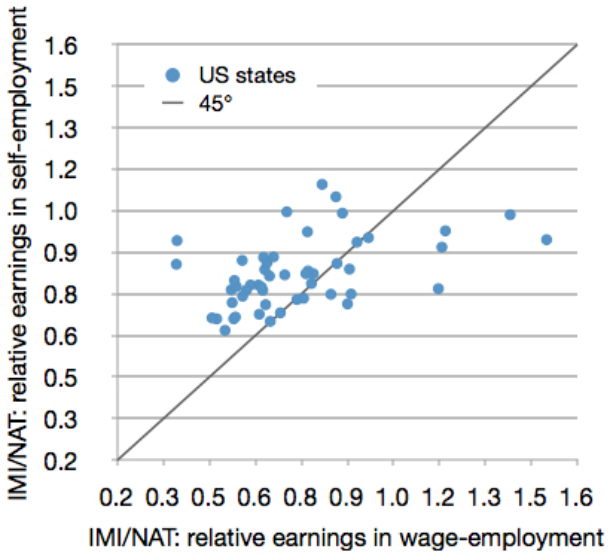


As relatively low earnings of immigrants are features of both discrimination and potential income gains to self-employment, these two effects may enforce each other in the development of immigrant entrepreneurship. Thus this chapter will conceptually build upon the work of the previous empirical analysis.

As far as empirical predictions are concerned, Figure 5.2 depicts the relation of the right and left-hand side of the above inequality across US states. It shows that in a vast majority of US states immigrants have a relative earnings advantage over natives in self-employment. Thus it seems that if they are discriminated on the labor market, self-employment should be a feasible alternative from the theoretical point of view[1].

[1]Due to methodological constraints the analysis conducted in this chapter will comprise only a selection of states. Neither of them are the potential outliers positioned far to the right - Kentucky, Maine, North Dakota, West Virginia, Wyoming - or left - District of Columbia, South Dakota - of Figure 5.2.

Figure 5.2: Immigrants employment share and earnings, by state and form of labor market activity



Source: own calculations based on American Community Survey 5-year PUMS, 2005-2010.

Given the financial benefits of pursuing self-employment rather than wage-employment and the detrimental effects labor market discrimination, this chapter poses the following hypotheses:

Hypothesis 5 *Labor market discrimination is related to increased intensity of immigrant self-employment.*

In testing the above hypothesis this thesis considers two issues, which are often neglected in the literature, but which seem more than important for the proper identification and interpretation of the research. The first issue is related to the outcome variably, i.e. immigrant self-employment. The second is related to discrimination as a notion which implies a relative position of one group against another.

As far as the outcome variable of this study is concerned, the literature often formulates hypotheses concerning the *propensity* to become self-employed (see e.g.: Borjas and Bronars, 1989; Clark and Drinkwater, 2000). This chapter proposes that the *intensity* of self-employment be considered instead. The reason for such a formulation is that when using the extent of discrimination as a predictor variable in estimating the probability of self-employment incidence we do not consider actual switchers. Thus it is impossible to determine whether individuals affected by discrimination become more inclined toward self-employment. All we can actually say without access to panel data is that the greater the extent of discrimination, the greater the share of self-employed in the labor force.

As far as the second issue is concerned, the question is which group should the outcomes of the group if interest be compared to. Previous research in the field seems to take it for granted that the comparison group is the "majority". In the case of the UK "whites"

are assumed to be such a group (see e.g. Clark and Drinkwater, 1998, 2000), "second generation Australians" in the context of Australia (see e.g. Evans and Kelley, 1991), and in the context of the US. "Non-Hispanic whites" or "British whites" (see e.g. Borjas and Bronars, 1989; Fairlie and Meyer, 1996). Evans and Kelley (1991) have noticed that discrimination can also be a measure of social distance, though.

The greater the ethno-cultural differences between the dominant and the minority group, the bigger the extent of discrimination may be. Of course this will be the case if employers also represent the dominant group. Yet this must not necessarily be the true. The notion that minority employers may actually discriminate against majority employees and favor their own have been extensively explored within the discussion on enclave economies (Sanders and Nee, 1987; Portes and Jensen, 1989, 1992; Sanders and Nee, 1992; Portes and Zhou, 1996). On the other hand, greater social distance may imply that a very distinct group is not the group of reference for a given immigrant minority. If this was the case then we could expect that even in case of large income differences, they would not be meaningful enough to affect an individual's labor market behavior. The arguments on the significance of various groups of reference in the migration process have been raised in the New Economic of Labor Migration (Stark and Bloom, 1985, see Chapter 2 for a review). This study will thus contribute to previous work by looking at various possible reference groups, rather than just one. In this sense it will consider a number of discrimination measures and evaluate them with respect to their association with the probability of becoming self-employed.

The main contributions of this study will be, first, providing a refined analysis of immigrant labor market discrimination, going beyond a simple comparison to one ethnic group. Second, demonstrating how the extent of labor market discrimination in relation to various groups of reference is associated with the intensity of immigrant self-employment. Third, in pursuit of these objectives this chapter will employ techniques well known from the field of labor economics, but which have been overlooked in some of the previous analysis leading to possibly biased results.

The chapter is structured as follows: the next section presents a review of research strategies related to the measurement of wage gaps concerning immigrant populations. The following section presents the methodology. It describes the data, sets the context and explains the choice of comparison groups for the wage gap analyses, and introduces the econometric models used in the analyses. The latter part is divided into two parts, the first explaining the technique used for determining the extent of discrimination, and the second explaining the technique used for quantifying the relationship between the extent of discrimination and the intensity of self-employment. Next, the results are presented and discussed, followed by a concise description of a external validity study. The summary and conclusions close the chapter.

5.2 Review of research strategies

Recent analyses on immigrant-native wage gaps prove the existence of significant differences in the earnings of non-natives and autochthons, both in the case of the US and other countries. How these gaps affect self-employment choices of immigrants is a less researched topic. Among the few written works on this issue Fairlie and Meyer (1996) explore self-employment rates of different ethnic groups in the US. Relaying on the works of Bonacich (1973), Light (1972), and Borjas (1986) they find that: the average differences between wage-employment and self-employment earnings is positively related to the self-employment rate of a given group; the more advantaged groups in terms of earned or unearned income have higher self-employment rates; and that higher self-employment rates in the home countries of immigrants do not necessarily translate into high self-employment rates at the destination.

In evaluating the discrimination hypothesis Fairlie and Meyer (1996) propose a two step procedure. They first estimate regressions where the dependent variables are wage-employment and self-employment earnings, and unearned income, and the independent variables are a number of individual characteristics, including dummies for ethnicity. In the second step they use the coefficients of the ethnicity dummies as independent variables in selectivity corrected probit models of self-employment. By means of this method they find that the discrimination theory does not hold, i.e. that ethnic penalty, *ceteris paribus*, does not lead to higher self-employment rates, but the contrary. The weakness of Fairlie and Meyer's (1996) approach is that by using dummy variables as controls for ethnicity, they provide biased estimators. Such a strategy is inadequate for measuring the extent of discrimination as it ignores potential differences in the other explanatory variables between the compared groups.

Clark and Drinkwater (1998) explore differences in wages between non-whites and whites in the UK based on the theoretical model of Coate and Tennyson (1992). They compute a non-white discrimination coefficient, which is the natural exponent of the β-coefficient on non-whites in a log-wage equation. Their findings point to a significant earnings disadvantage of non-whites, both in payed- and self-employment. Such a strategy suffers from the same drawbacks as that of Fairlie and Meyer's (1996), though. In subsequent work Clark and Drinkwater (2000) take the difference in the expected earnings from self- and wage-employment as an explanatory variable in estimating the probability of becoming self-employed. Their findings confirm that self-employment intensity is positively related to the disparity in one's predicted earnings from the two forms of labor market activity.

In 2001 Nielsen et al. took advantage of the Oaxaca-Blinder decomposition and proposed a crucial methodological refinement in decomposing wage differences between the immigrants and natives, which took into consideration not only the discrimination hypothesis, but also assimilation theory. Thy estimate the wage gaps between immigrants and natives based on Danish data disaggregating the differences into three components (Nielsen et al., 2001, p.10)

$$\Delta_E = \hat{y}^n - \hat{y}^i = \underbrace{\hat{y}^n - \hat{y}^{i*n}}_{Q_E^*} + \underbrace{\hat{y}^{i*n} - \hat{y}^{i*i}}_{D_E^*} + \underbrace{\hat{y}^{i*i} - \hat{y}^i}_{A^*}$$

where Q_E^* is the difference between the native (n) and immigrant (i) earnings when evaluated at natives' coefficients, i.e. it indicates the part of the wage gap resulting from differences in endowments between the two groups; D_E^* is the difference between the native and immigrant earnings which is related to differences in coefficients, rather than endowments, i.e. it indicates the extent of discrimination; and finally A^* is the part of the wage gap which is due to the fact that the immigrant-native wage gap may decrease as immigrants assimilate (have longer duration of stay and longer working experience in Denmark). The analysis considers two hypothetical states (*). "Perfect assimilation" is after 10 years of stay and 10 years of work experience. The "unassimilated state" concerns newcomers. The two states are operationalized by means of setting relevant variables such as years since migration, experience and age to values corresponding to perfect assimilation and being a newcomer. Nielsen et al. (2001) do not find strong evidence for wage discrimination against immigrants, but they do find large differences in their work experience (shorter employment spans than of the native population)[2].

In a cross country study of 17 European economies van Tubergen (2005) uses the native and immigrant employment rates, rather than wage-gaps, to measure the extent of discrimination. The findings support his hypothesis that greater probability of becoming self-employed is associated with lower levels of employment in a given ethnic community. The imperfection of this measure lies in the various ethnicity-specific features unrelated to discrimination, which affect labor market choices and are captured by the employment rate (e.g. social roles within households, time use arrangements, fertility patterns). An analog analysis to that of van Tubergen (2005) has been done for wage gaps between immigrants and natives across different ethnic groups in Sweden (Hammarstedt, 2006).

Most recently Lehmer and Ludsteck (2011) described the immigrant-native wage gaps for Germany. They specifically focused on the period 1995-2000. Oaxaca-Blinder decompositions (ref. Nielsen et al., 2001) of earnings revealed greatest labor market discrimination among Poles, yet this pattern was not applicable to all East European citizens. This, similarly to studies such as that of Fairlie and Meyer (1996), points to the fact that labor market mechanisms affecting "aggregate ethnic groups", or a certain selection of ethnic groups on average, may in fact work quite differently for specific ethnicities. Interestingly Lehmer and Ludsteck (2011) propose to distinguish between "pure wage discrimination" and discrimination resulting from pushing immigrants into lower payed jobs, i.e. "occupational discrimination". In order to do so they run wage decompositions both controlling for sector of employment and not.

From the reviewed literature it follows that, first, focusing on one ethnic group rather than an aggregate may be very meaningful. Second, it shows that who is treated as

[2]The authors do not conduct further analysis on how these wage-gaps affect immigrants' intensity of self-employment, though.

the "dominant" comparison group is quite arbitrary and varies across studies. Third, we can see that the Oaxaca-Blinder wage decomposition method and its variations has been successfully applied to immigrant-native wage comparisons. Last, but for this thesis definitely not least, it is clear that despite various attempts at measuring the extent of discrimination it has less frequently been analytically considered as a stimulant for self-employment.

5.3 Methodology

This analysis has two components. The first inquiry deals with estimating the extent of labor market (wage) discrimination between Puerto Rican immigrants to the US and the possible reference groups of natives. The second inquiry looks at how the extent of wage discrimination affects the probability of becoming self- employed.

5.3.1 Data

In this study we use the 5-year American Community Survey data from the year 2005-2010 (Ruggles et al., 2010). It is the same dataset as has been used for the analysis in Chapter 4. In contrast to that chapter, though, for this part of the research the Puerto Rican Community Survey data is not used, as here the focus is only on the US as a receiving country. The descriptive statistics for selected variables used in the analyses are presented in Table 5.1.

Table 5.1: Means and standard deviations of selected variables used

Variable	Puerto Rican immigrants	Puerto Rican natives	Hispanic natives	Non-Hispanic white natives	white natives
age	45.36	33.94	34.43	43.97	43.58
	(14.58)	(12.68)	(14.49)	(15.21)	(15.30)
gender-male	0.47	0.48	0.49	0.49	0.49
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
gender-female	0.53	0.52	0.51	0.51	0.51
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
married-no	0.50	0.64	0.60	0.41	0.42
	(0.50)	(0.48)	(0.49)	(0.49)	(0.49)
married-yes	0.50	0.36	0.40	0.59	0.58
	(0.50)	(0.48)	(0.49)	(0.49)	(0.49)
education-incomplete	0.03	0.01	0.01	0.00	0.00
	(0.16)	(0.08)	(0.08)	(0.06)	(0.06)
education-primary	0.11	0.03	0.04	0.01	0.01
	(0.31)	(0.16)	(0.18)	(0.12)	(0.12)
education-secondary	0.51	0.58	0.58	0.46	0.46
	(0.50)	(0.49)	(0.49)	(0.50)	(0.50)
education-tertiary	0.36	0.39	0.38	0.53	0.52
	(0.48)	(0.49)	(0.49)	(0.50)	(0.50)
agriculture	0.02	0.01	0.02	0.03	0.03
	(0.12)	(0.10)	(0.14)	(0.16)	(0.16)
construction	0.05	0.05	0.07	0.07	0.07
	(0.22)	(0.22)	(0.25)	(0.26)	(0.26)
manufacturing	0.12	0.07	0.08	0.12	0.12
	(0.33)	(0.26)	(0.28)	(0.32)	(0.32)
transport/communications	0.08	0.08	0.07	0.06	0.06
	(0.26)	(0.27)	(0.25)	(0.25)	(0.25)
wholesale	0.03	0.03	0.03	0.03	0.03
	(0.17)	(0.16)	(0.17)	(0.17)	(0.17)
retail	0.16	0.21	0.22	0.17	0.17
	(0.36)	(0.41)	(0.41)	(0.37)	(0.38)
finance/insurance	0.06	0.07	0.06	0.07	0.07
	(0.25)	(0.26)	(0.24)	(0.25)	(0.25)
business/repair services	0.07	0.07	0.06	0.06	0.06
	(0.26)	(0.25)	(0.24)	(0.23)	(0.23)
personal services	0.04	0.04	0.03	0.03	0.03
	(0.20)	(0.19)	(0.18)	(0.17)	(0.17)
entertainment	0.02	0.02	0.02	0.02	0.02
	(0.12)	(0.14)	(0.14)	(0.13)	(0.13)
professional services	0.28	0.28	0.26	0.29	0.29
	(0.45)	(0.45)	(0.44)	(0.45)	(0.45)
public administration	0.06	0.06	0.05	0.05	0.05
	(0.23)	(0.23)	(0.23)	(0.21)	(0.21)
manager, professional	0.22	0.22	0.22	0.32	0.32
	(0.42)	(0.42)	(0.41)	(0.47)	(0.47)
technician, salesman, etc.	0.27	0.36	0.35	0.30	0.30
	(0.45)	(0.48)	(0.48)	(0.46)	(0.46)
service worker	0.21	0.20	0.19	0.14	0.14
	(0.41)	(0.40)	(0.34)	(0.35)	(0.35)
farmer etc.	0.02	0.01	0.02	0.03	0.03
	(0.13)	(0.11)	(0.14)	(0.16)	(0.16)
precision or repair worker, etc.	0.09	0.08	0.09	0.10	0.10
	(0.29)	(0.26)	(0.29)	(0.30)	(0.30)
operator, fabricator, laborer	0.17	0.11	0.12	0.11	0.11
	(0.37)	(0.31)	(0.33)	(0.31)	(0.31)
Speaks English-none	0.04	0.00	0.00	0.00	0.00
	(0.20)	(0.04)	(0.06)	(0.01)	(0.02)
Speaks English-only	0.11	0.39	0.41	0.97	0.95
	(0.31)	(0.49)	(0.49)	(0.17)	(0.23)
Speaks English-very well	0.50	0.53	0.50	0.02	0.05
	(0.50)	(0.50)	(0.50)	(0.16)	(0.21)
Speaks English-well	0.21	0.06	0.07	0.00	0.01
	(0.41)	(0.23)	(0.25)	(0.06)	(0.08)
Speaks English-not well	0.13	0.02	0.02	0.00	0.00
	(0.34)	(0.12)	(0.13)	(0.04)	(0.05)
working-yes	0.72	0.80	0.79	0.84	0.84
	(0.45)	(0.40)	(0.41)	(0.36)	(0.37)
self-employed-yes	0.06	0.05	0.06	0.11	0.11
	(0.24)	(0.21)	(0.24)	(0.32)	(0.31)
Observations	49,031	51,823	532,706	7,129,412	7,464,142

Source: The IPUMS databases, Ruggles et al. (2010). Standard deviations in parenthesis.

Relatively many Puerto Rican immigrants in the US have primary education in comparison to the other groups. In contrast, there are relatively fewer Puerto Rican immigrants with tertiary education, especially when compared to white and non-Hispanic white natives. In terms of industrial clustering there are no striking differences. Puerto Rican immigrants are more often present in manufacturing and less present in retail when compared to Puerto Rican natives or Hispanics in general, but their shares in these two industries are similar to those of whites and non-Hispanic whites. As far as occupations are concerned Puerto Rican immigrants have a visibly higher presence as operators, fabricators and laborers. Differences are also visible in terms of language proficiency of Puerto Rican immigrants, who rather rarely speak only English, and more often (than the other groups) declare they speak English "well" or "not well". The Puerto Rican immigrant self-employment rate is also lower than that of whites and non-Hispanic whites. It is comparable to that of Puerto Rican and Hispanic natives, though.

5.3.2 Empirical setting

In light of the considerations presented in the introduction to this chapter, taking the dominant group as a comparison and assuming that most employers also come from that group, thus capturing the "maximal" discrimination, seems justified from the perspective of labor market analysis. When considering discrimination as a factor affecting individuals' labor market decisions, though, the choice of reference group is not as obvious. Though the discrimination of, for example, Ugandan immigrants in the US when compared to white US citizens may be largest in numerical terms, it may not actually push the former out of wage-employment into self-employment. Instead numerically smaller, but "socially" more significant discrimination may play a critical role, e.g. discrimination of Ugandans in comparison to the black population, to black immigrants, or specifically to black immigrants from East African countries.

This chapter, continuing to analyze Puerto Rican immigrants to the US, will thus consider several ethnic groups of natives for comparisons:

- **Puerto Rican natives**, i.e. US born persons of Puerto Rican ancestry;
- **Hispanic natives**, i.e. US born persons of Hispanic, including Puerto Rican, ancestry;
- **White natives**, i.e. white US born persons (note that all Hispanics, including Puerto Ricans are classified as white);
- **White non-Hispanic natives**, i.e. white US born persons, excluding all Hispanics.

The rates of Puerto Rican immigrant self-employment in comparison to the selected groups of natives are presented in Table 5.2.

The fact that the self-employment rate of Puerto Rican immigrants in the US is higher than that of Puerto Rican and, generally, Hispanic natives, but lower than that of non-Hispanic whites and white natives creates an interesting context for the analysis. If self-

Table 5.2: Self-employment rates by selected groups

	self-employment	wage-employment
Puerto Rican immigrants	6.23%	93.77%
Puerto Rican natives	4.52%	95.48%
Hispanic natives	6.01%	93.99%
All white non-Hispanic natives	11.22%	88.78%
All white natives	11.03%	88.97%

Data: Ruggles et al. (2010).

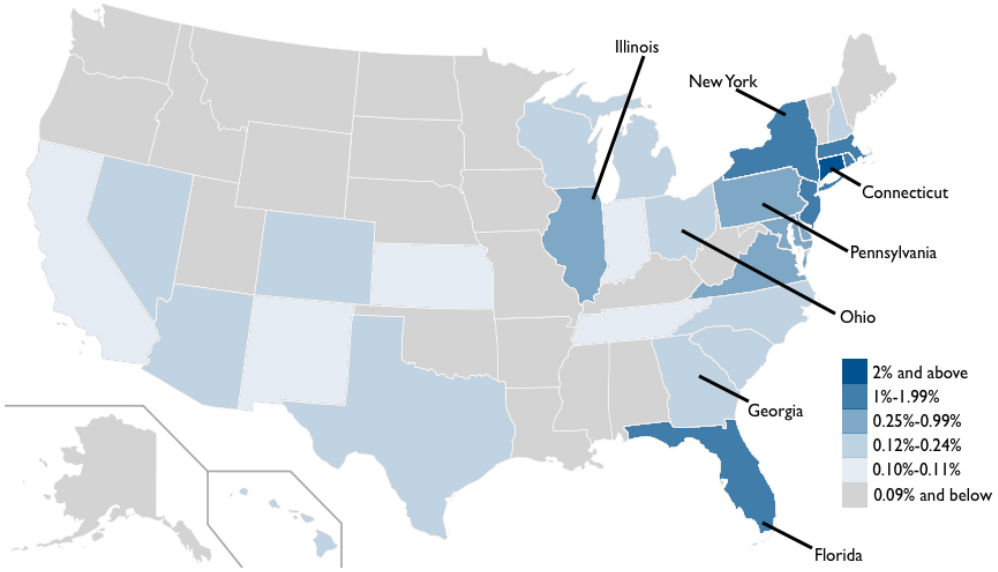
employment is in fact driven by the extent of labor market discrimination should we be also expecting higher discrimination of Puerto Rican immigrants when compared to Puerto Rican and Hispanic natives, but lower discrimination when compared to non-Hispanic white natives or white natives in general? This type of analysis will be a much needed addition to the discrimination literature, which usually considered the diversity of immigrant groups, rather than of natives (for an analysis of discrimination of various US immigrant groups see: Fairlie and Meyer, 1996).

As this analysis considers discrimination and self-employment intensity of one group only, i.e. Puerto Rican immigrants to the US, the regional diversity of labor markets will be exploited to achieve variation in the effect of discrimination on self-employment outcomes. The state level has been chosen as the basis for regional comparisons. On one hand this is due to state-level specificity concerning immigrant incorporation and the labor market in general. On the other hand only such a level of analysis allows to explore solely Puerto Ricans as an ethnic minority, assuring large enough numbers of observations. Due to the relatively small number of Puerto Ricans in general, and consequently of the Puerto Rican immigrant population in the US, and because of their clustering in certain states, only some areas have been included in final the analyses, though. Figure 5.3 presents the density of Puerto Rican immigrants relative to the total population across US states. The seven states eventually included in the analyses have been labeled.

The state with the highest share of Puerto Rican immigrants is Connecticut, with over 2% of Puerto Rican immigrants in the total population. It is a relatively small state, though, with 3.5 million inhabitants. Florida's population, despite being comprised "only" in 1.94% of Puerto Rico born persons, has nearly 360 thousand Puerto Ricans living there, what makes it the state with the largest Puerto Rico born population.

The seven states which been included in the analysis of discrimination: Connecticut, Florida, Georgia, Illinois, New York, Ohio, and Pennsylvania comprise 65% of the Puerto Rican immigrant population in the US. In the remaining states computations of the extent of discrimination were not possible for one or more of the comparison groups. Due to the resulting incomparability of results across different samples these states were left out of the study.

Figure 5.3: Share of Puerto Rican immigrants across states



Source: Own elaboration based on map from Wikimedia Commons.

5.3.3 Estimating the extent discrimination

First, we measure to what extent are Puerto Rican immigrants discriminated in the labor market against different reference groups. As it has been indicated in the introduction to this chapter, the choice of several groups of natives for comparison was dictated by the need to consider the extent of discrimination as a decision-making factor in subsequent analyses.

Following the works of Nielsen et al. (2001) and Lehmer and Ludsteck (2011) the applied measure of discrimination is derived from the Oaxaca-Blinder decomposition of wage equations for specific subpopulations. The difference in expected earnings of two groups within these subpopulations - Puerto Rican immigrants and a native comparison group active on the US labor market - are computed from Mincerian wage equations and are Heckman-corrected for selection into working (Heckman, 1979):

$$Pr(working) = \beta_0 + \beta_1 \mathbf{X} + \beta_2 martial\ status + \epsilon, \tag{5.1a}$$

$$ln(earnings) = \beta_0 + \beta_1 \mathbf{X} + \beta_2 industry + \beta_3 occupation + \\ + \beta_4 \text{ Inverse Mill's Ratio} + \epsilon_i \tag{5.1b}$$

where $\mathbf{X}$ is a set of individual characteristics such as: age, gender, level of educational attainment, and English language proficiency; *industry* is the industry of one's activity[3]; *occupation* is one's occupation[4]; (5.1a) estimates the probability that an individual in

[3]Constructed based on a three-digit classification of industries coded in the IND1990 variable in the PUMS datasets, yielding 13 sectors.

[4]Based on a 3-digit classification of occupations derived from the OCC1990 variable in the PUMS datasets, 7 categories.

productive age (16-65) is actually working, and (5.1b) is the log-wage equation which corrects for that fact using the Inverse Mill's Ratio as a regressor. Marital status has been used as the selection variable which is assumed to affect the odds of working, but not the productivity. This set of regressions is run separately for each of the five considered groups (Puerto Rican immigrants, Puerto Rican natives, Hispanic natives, white natives and white non-Hispanic natives) for each state. Following Jann (2008) we then compute the differences in expected earnings, estimated from (5.1b), between Puerto Rican immigrants (PRi) and each of the comparison groups (CG) in each state.

The difference is then decomposed into three components

$$R = E(Y_{CG}) - E(Y_{PRi}) \implies \tag{5.2a}$$

$$\begin{aligned} \implies R = &[E(\mathbf{X_{CG}}) - E(\mathbf{X_{PRi}})]'\beta_{PRi} + \\ &+ E(\mathbf{X_{PRi}})'(\beta_{CG} - \beta_{PRi}) + \\ &+ [E(\mathbf{X_{CG}}) - E(\mathbf{X_{PRi}})]'(\beta_{CG} - \beta_{PRi}) \end{aligned} \tag{5.2b}$$

We are especially interested in $[E(\mathbf{X_{PRi}})'(\beta_{CG} - \beta_{PRi})]$ which measures the contribution of differences in the wage regression coefficients to the estimated wage inequality. The differences in coefficients are calculated from the view point of Puerto Rican immigrants, i.e. they are weighed by their levels of predictor variables. The measure quantifies the difference in earnings which would prevail even if Puerto Rican immigrants had the other group's characteristics. Considering this component of the wage difference, rather than the rest of (5.2b) allows us to filter out that part of the wage difference between Puerto Rican immigrants and a reference group which is due to differences in endowments, i.e. differences in levels of productivity-related measures. The remainder of the difference is then attributable to discrimination.

5.3.4 Estimating the relation of discrimination to self-employment intensity

As our interest lies in describing how the extent of wage discrimination affects the probability of becoming self-employed ($PR(se)$) among immigrants, a probit model will be utilized in the second stage of inquiry. The regression once more takes advantage of the Heckman-correction and the Inverse Mill's Ratio estimated in (5.1a) and the model has the following form

$$Pr(se) = \beta_0 + \beta_1 \mathbf{X} + \beta_2 \text{discrimination} + \beta_3 \text{Inverse Mill's Ratio} + \epsilon \tag{5.3}$$

The measure of discrimination inputted into (5.3) is the second summand of (5.2b) divided by the value of earnings, i.e.

$$\text{discrimination} = \frac{[E(\mathbf{X_{PRi}})'(\beta_{CG} - \beta_{PRi})]}{\overline{Y}} \tag{5.4}$$

Given that earnings are always positive, this measure takes the sign of the discrimination ($-$ for positive discrimination, $+$ otherwise) and is standardized by the mean earnings in a given state ($\overline{Y}$). This standardization allows to assess the relative extent of discrimination. This procedure was applied for two reasons. First, it is critical to know whether we are talking about positive discrimination, i.e. preferential treatment in terms of remuneration, or "regular" discrimination. And second it is crucial to know not only how many dollars more (or less) would one earn if it was not for discrimination, but also what part of the actual wage would that make up for. In case of USD 12,000 earnings per year a USD 1000 difference would be worth a month's pay, but if the wage is USD 36,000, then that only makes it worth 10 days' pay, being relatively less discriminatory. The average state wage was used in order to limit possible endogeneity between the probability of self-employment and the income measure.

5.4 The role of discrimination in immigrant self-employment

The average level of discrimination, as defined by (5.2b), of Puerto Rican immigrants when compared to the different groups considered in this chapter yields results presented in Table 5.3. The percentage values indicate how much less (or more) do Puerto Rican immigrants earn due to labor market discrimination. Given that the estimated wage-equations were regressions of log-earnings, the percentage differences in earnings were calculated using the following formula

$$\%value = \{(exp[E(\mathbf{X_{PRi}})'(\beta_{CG} - \beta_{PRi})])^{-1} - 1\} \times 100\% \tag{5.5}$$

Taking the inverse allows to consider the earnings of natives (any of the comparison groups) as 100% and relate the earnings of immigrants to that value. Subtracting 1 implies that the presented results show by how many percent the value of estimated immigrant wages is above or below the natives' wages. Notably, these values do not indicate the actual difference in wages, but only the difference which is attributable to discrimination. In fact, the difference might be relatively larger (smaller) due to poorer (better) endowment in productivity-related characteristics of Puerto Rican immigrants. Following Lehmer and Ludsteck (2011) the wage decompositions have also been run without accounting for occupations. Analyses of discrimination in terms of hourly earnings, both including and excluding occupations, have also been performed. The results slightly differed in magnitude, but never in terms of sign. Thus here the focus will be on the results presented in terms of yearly earnings and inclusive of occupations. Tables with the results of the other specifications are presented in the Appendix.

Though the extent of discrimination greatly differs across states and comparison groups, from table 5.3 it follows that in case of comparisons against Hispanic and Puerto Rican natives Puerto Rican immigrants are, on average, negatively discriminated. On the state level this regularity is visible only in the state of New York.

Table 5.3: Labor market discrimination of Puerto Rican immigrants across states and comparison groups

	Puerto Rican natives	Hispanic natives	white natives	Non-Hispanic white natives
Connecticut	5%	-3%	29%	28%
Florida	-22%	-29%	-10%	-7%
Georgia	-19%	-24%	-5%	-2%
Illinois	20%	4%	42%	10%
New York	-4%	-1%	66%	69%
Ohio	37%	-12%	12%	12%
Pennsylvania	80%	0%	53%	56%
average effect	negative	negative	positive	positive
specification of discrimination measures				
occupations	+	+	+	+
earnings	yearly	yearly	yearly	yearly

Only Oaxaca-Blinder decomposition discrimination coefficients significant below the 5% level have been considered. The discrimination measure took the value of 0 otherwise. Only states for which models achieved convergence are shown.

Only in two states Puerto Rican immigrants enjoy an earnings advantage in comparison to all four reference groups - in Illinois and Pennsylvania. The latter state also represents the most striking positive discrimination. Puerto Rican immigrants living in Pennsylvania earn 80% more than Puerto Rican natives. Positive discrimination of a relatively large magnitude has been also recognized in New York, yet there an income gain of 66% and 69% is in comparison to white and non-Hispanic white natives, respectively.

When it comes to negative discrimination the highest values were recognized in Florida. Puerto Rican immigrants turned out to be discriminated when compared to all native groups. In Florida Puerto Rican immigrants also experienced the highest level of discrimination. Their wages are estimated to be 29% lower than those of Hispanic natives living in the state. Negative discrimination is also consistently observed across all comparison groups in Georgia.

The patterns represented by the Oaxaca-Blinder decompositions open a new array of research questions. It appears that the institution of labor market discrimination is far more complex than expected. For one thing it varies across localities implying high context-specificity. Second, it differs greatly across comparison groups. It seems that the more general (distant) the group, the more favorable working conditions one may obtain. This finding contradicts the intuition that the more distant the group, the greater the discrimination would be.

These patterns are explainable under one scenario, already signaled in the introduction to this chapter, i.e. under the circumstance that the employers are in fact from the minority group. This would explain favoring of Puerto Rican immigrants as a form of within-group support. The fact that in some cases, mostly when compared to Puerto Rican or Hispanic natives, Puerto Rican immigrants are negatively discriminated might be indicative of an

ethnic market in which labor market institutions work as expected towards insiders, but when compared to outsiders, labor market favoring dominates.

Given the obtained measure(s) of labor market discrimination we have moved to estimating the relation of discrimination to the probability of becoming self-employed. Marginal effects of selected independent variables after estimating (5.3) are presented in Table 5.4

Table 5.4: Probit model output, marginal effects

	base case	Puerto Rican natives	Hispanic natives	white natives	Non-Hispanic white natives
discrimination		0.438***	0.672***	0.363***	0.354***
age	0.008***	0.008***	0.009***	0.009***	0.009***
gender (1=female)	-0.016***	-0.016***	-0.018***	-0.018***	-0.018***
education			incomplete - reference category		
primary	-0.018***	-0.016***	-0.016***	-0.016***	-0.016***
secondary	-0.018***	-0.016***	-0.015***	-0.016***	-0.016***
tertiary	-0.003	-0.004	-0.003	-0.003	-0.003
N	19,582	19,582	19,582	19,582	19,582
English proficiency	+	+	+	+	+
occupations	+	+	+	+	+
industries	+	+	+	+	+
specification of discrimination measures					
occupations	n/a	+	+	+	+
earnings	n/a	yearly	yearly	yearly	yearly

* significant at 10% level, ** significant at 5% level, *** significant at 1% level

In each of the considered cases discrimination has a positive marginal effect in relation to the probability of self-employment. Thus it seems that we can confirm the initial hypothesis that the larger the discrimination, the higher the probability of becoming self-employed.

From Table 5.4 it follows that labor market discrimination of Puerto Rican immigrants when compared to Hispanic natives has the largest relation to the probability of Puerto Rican immigrant self-employment. If Puerto Rican immigrants compare their wages to non-Hispanic white natives, then their intensity of self-employment is relatively the smallest. If white or Puerto Rican natives are the reference group, the relation of wage discrimination to the probability of becoming self-employed falls somewhere in the middle.

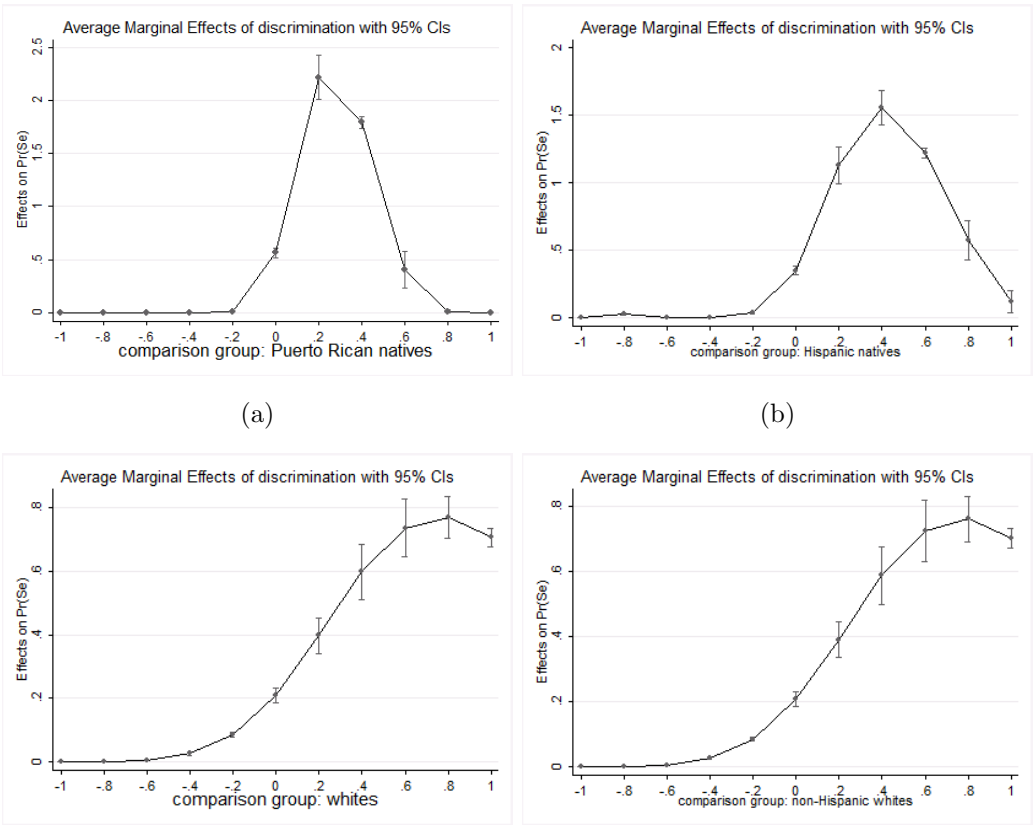
These results tell the story of social distance, as proposed by Evans and Kelley (1991). Whites and non-Hispanic whites are the two groups furtherest in social distance from Puerto Rican immigrants. Whites are inclusive of Puerto Ricans, but comprising many other ethnic groups. Non-Hispanic whites exclude Puerto Ricans all together. Thus the fact that the wage comparison to these two groups exhibits the weakest relation to the intensity of self-employment in a given state is in line with the theory.

Puerto Rican natives were considered the group which is the nearest to Puerto Rican immigrants in terms of social distance. The relation of discrimination to the intensity of self-employment is in this case stronger than that in the case of whites and non-Hispanic

whites, but weaker when compared to Hispanic natives in general. What is important to recall is that in case of comparisons with native Puerto Ricans and Hispanics average discrimination across the seven considered states is positive. Thus the results from these two cases actually show that the larger the favoring (the lower the discrimination) of Puerto Rican immigrants, the smaller the intensity of self-employment. It seams reasonable to assume that smaller favoring in comparison to a similar group would have a weaker effect on the intensity of self-employment. In case of a very similar group equality in terms of earnings could be very much acceptable. In such a context smaller wage advantages must not necessarily be associated with higher probability of self-employment.

In order to gain better understanding of the obtained results we consider a specific case, i.e. an average aged male, working in the retail sector, in a technical, sales, administrative position, speaking English very well, with secondary education. The marginal effects of various discrimination measures on the predicted probability of self-employment calculated for such a person are presented in Figure 5.4[5].

Figure 5.4: Marginal effects of various levels of discrimination on the Probability of Puerto Rican immigrant self-employment, by comparison group



From Figure 5.4 we can see that for positive discrimination (negative value of our discrimination measure, left part of the OX axes) its relation to the probability of becoming

[5]The figures consider discrimination measured with respect to occupations and yearly earnings. Results for analyses where earnings are hourly and discrimination does not account for occupations are presented in the Appendix.

self-employed is very small. When negative discrimination is in place in most cases it doubtlessly increases the probability to become self-employed, as predicted by the theories of immigrant self-employment. In case of discrimination of Puerto Rican immigrants when compare to Puerto Rican natives (a) or Hispanics (b) the relation is not as obvious. At very high values of discrimination the probability of immigrant self-employment seems to fall.

Table 5.5 provides a precise numerical value of the probability of becoming self-employed when the marginal effects of the various measures of discrimination are evaluated at their mean values and mean (dominant, where applicable) values (categories, where applicable) of independent variables[6].

Table 5.5: Marginal effects of various discrimination measures evaluated at their mean on the probability of Puerto Rican immigrant self-employment, by comparison groups

	dy/dx	Std. Err.	z	P>z	[95% Conf.	Interval]
Puerto Rican natives	0.59	0.02	24.29	0.00	0.55	0.64
Hispanic natives	0.40	0.02	22.00	0.00	0.36	0.43
white natives	0.19	0.01	19.33	0.00	0.17	0.21
white non-Hispanic natives	0.19	0.01	18.69	0.00	0.17	0.21

Under such a scenario the relation of discrimination to the probability of elf-employment seems linear in relation to social distance. The greater the distance, the smaller the interdependence between wage discrimination and self-employment intensity.

5.5 External validity of the study

Similarly to what has been done in the case of the analysis of income gains to immigrant self-employment, this chapter also provides an extension of the above study. The subject of the following exercise are all immigrants to the US. Their earnings are compared to the earnings of their co-ethnics who are US citizens. The immigrant and native populations comprise seven ethnic groups delimited by approximately continental divisions: West Europeans, North Africans and Southwest Asians, Hispanics, Central Europeans, Sub-Saharan Africans, Asians, and North Americans[7]. All models were computed on the 2005-2010 American Community Survey dataset (Ruggles et al., 2010). The Oaxaca-Blinder decomposition of wage differences is done for citizens and non-citizens within these ethnic groups.

An attempt was also made to compute the extent of discrimination not only within ethnic groups, but identically to the strategy applied for Puerto Ricans, also across states. As the various ethnic groups have different concentrations across states, computing discrimination for all the ethnic groups was possible only for a handful of locations, California, Florida, and New York. The immigrant population in these three states amounts to ∼35%,

[6]The results of this exercise for other discrimination measures are presented in the Appendix.

[7]Defined based on the ANCESTR1 variable in the PUMS datasets. The variable captures the first response to a question on one's ethnic origin

unlike the high concentration (∼62%) of Puerto Ricans in the seven states selected for the main analysis (Connecticut, Florida, Georgia, Illinois, Ohio, New York and Pennsylvania). These states are also not necessarily the major receiving states for immigrants from the specific ethnic groups. Thus only the variability of average discrimination across ethnic groups was eventually considered.

In order to determine the expected sign of the relation of discrimination to the intensity of self-employment for all immigrants we have tested if self-employment is not an even less profitable opportunity than discriminatory wage employment. This could have been the case because, similarly to the situation on the wage-earning labor market, (ethnic) group membership may generate a signaling effect also in the case of the self-employed, resulting in relatively lower expected earnings and less incentive to choose self-employment as a labor market strategy. The mechanisms behind such outcomes could be consumer or credit market discrimination. If immigrants earned less in self-employment than in (discriminatory) wage-employment, then we would not expect them to become self-employed more often nevertheless.

In order to test for the possibility of lower earnings from self-employment an additional analysis using Propensity Score Matching has been applied (see Chapter 4 for methodological details). The method was identical to that used in Chapter 4, the only difference being that the control group were all wage-employed immigrants (non-citizens) while the treated group were all self-employed immigrants. In order to facilitate computations a random 10% sample has been selected from the 10 million observations in the 5-year ACS used for the discrimination analysis. Immigrants constituted 7.33% of the 1,014,340 observations eventually considered. 11% (59,665 observations) of the working immigrant population are the self-employed. The estimation of propensity scores was done on 51,285 observations, all of which were on common support.

Kernel matching, bootstrapped estimates of treatment effects obtained after kernel matching, and one-to-one matching have indicated no significant difference between the estimated average earnings (yearly and hourly) of immigrant wage-employed and self-employed workers. All matching procedures assured acceptably similar samples after matching (bias of variables used for propensity score estimation lower than 5% between the treated and control group).

The fact that immigrants' earnings from wage-employment and self-employment do not differ does not yet imply that immigrants are on average indifferent between the two forms of labor market activity. Recalling the theoretical models we have to note that what maters is the difference between the relative earnings of self-employed and wage-employed immigrants to respective groups of natives. The theory predicts that if discrimination persists, minority members will opt for self-employment until their crowding-in pushes profits low enough that

$$w_M/w_N = \pi_M/\pi_N$$

where w denotes wages, π denotes profits from entrepreneurship and indices M and N indicate minority and non-minority members, respectively (or migrants and natives, in this case).

Knowing that the enumerators in the above equation are statistically identical ($w_M = \pi_M$), the difference between the denominators is what remains unexplored. Thus Propensity Score Matching has been applied once more, this time to compare the difference in earnings between self-employed and wage-employed natives. The methodology was identical to what has been done heretofore. The 10% random sample of the 5-year ACS has been used, just as for the comparison between immigrant wage-employment and self-employment earnings. 635,581 observations were eventually included in the analysis, all on common support. The self-employed constituted 13% of the sample[8]. Both kernel and one-to-one matching provided satisfactory comparability between the characteristics of wage-employed natives (the control group) and self-employed natives (the treated group). The comparison of outcomes (hourly and yearly earnings) conducted for samples matched by means of both techniques yielded the same results presented in Table 5.6. Due to obtaining results which are highly significant and robust to different matching procedures, bootstrapping has not been performed.

Table 5.6: Income gains to self-employment among US citizens: Average Treatment Effects on the Treated (ATT)

Variable	Treated	Controls	Difference	PSM S.E.	T-stat
Kernel Matching					
hourly income	25.93	21.47	4.46	0.98	4.55
yearly income	52,339	45,274	7,065	1,951	3.62

From Table 5.6 it follows that native self-employed persons earn on average significantly more, both in hourly and yearly terms, than wage-employed natives. This would indicate that since $w_M = \pi_M$ and $w_N < \pi_N$, then $w_M/w_N > \pi_M/\pi_N$, what, according to the theory, would be the reason why immigrants, in general, despite discrimination, are not inclined to choose self-employment over wage-employment. This result leads to hypothesizing that the relation between discrimination and the probability of self-employment among US immigrants in general should not be positive.

The Oaxaca-Blinder decomposition of wage differences between immigrants and natives within the seven aggregated ethnic groups gave the following results:

Table 5.7 indicates varying levels of discrimination. Non-citizens of Western European ancestry are the only group which has a negative value of the coefficient-related part of

[8]In this study the fraction of self-employed immigrants is 11% while self-employed natives 13% implying a 2 percentage point difference in favor of the natives. This contradicts the data presented in Figure 1.1a, where the US is an example of a country in which the immigrant self-employment rate exceeds that of the natives. The differences may be due to two reasons. First, the OECD data defines immigrants and natives based on place of birth rather than citizenship, as it is done here. Second, they consider the actual self-employment rate, that is the number of self-employed in the total labor force, while in this study only the wage-employed and self-employed workers are considered in the analysis, excluding e.g. the unemployed.

Table 5.7: Oaxaca-Blinder Decomposition of Wage Gaps by Ethnic Groups

	Wage equation		Decomposition		
	Coefficient	Observations	Endowments	Coefficients	Interaction
West European		2,941,563			
Predicted $ln(wage)$ for natives	10.18***				
Predicted $ln(wage)$ for immigrants	10.51***				
Difference	-0.33***		-0.25***	-0.09***	0.03***
North African & Southwest Asian		44,812			
Predicted $ln(wage)$ for natives	10.21***				
Predicted $ln(wage)$ for immigrants	9.94***				
Difference	0.27*		0.14***	0.13***	0.00
Hispanics		764,129			
Predicted $ln(wage)$ for natives	10.20***				
Predicted $ln(wage)$ for immigrants	9.69***				
Difference	0.51***		0.14***	0.38***	-0.01**
Central Europeans		356,433			
Predicted $ln(wage)$ for natives	10.48***				
Predicted $ln(wage)$ for immigrants	9.99***				
Difference	0.49***		0.32***	0.30***	-0.13***
Sub-Saharan African		44,933			
Predicted $ln(wage)$ for natives	10.27***				
Predicted $ln(wage)$ for immigrants	10.11***				
Difference	0.17***		0.13***	0.05***	-0.01
Asian		316,760			
Predicted $ln(wage)$ for natives	10.51***				
Predicted $ln(wage)$ for immigrants	10.15***				
Difference	0.36***		-0.01*	0.33***	0.04***
North American		1,360,243			
Predicted $ln(wage)$ for natives	10.26***				
Predicted $ln(wage)$ for immigrants	10.25***				
Difference	0.01		-0.22***	0.14***	0.09***

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

the decomposition. This means that they experience positive discrimination. That is to say that if they had the same characteristics as the natives and were evaluated at the values of β-coefficients at which they are now, they would earn less, than they do. Non-citizens from Western Europe are, based on the observables considered in the wage equation, also more productive, than citizens (negative value of the endowment-related part of the decomposition). West European immigrants are also the only subpopulation which earns significantly more than their native counterparts. The rest of the ethnic groups do experience significant, negative labor market discrimination and as immigrants they earn significantly less than the natives. The exception are people of North American origin (e.g. Afro-American, African-American,Central/North American Indian, Canadian, American, United States), who do not experience any significant difference in earnings whether they have US citizenship or not.

Among Asians discrimination is responsible for 91% of the wage gap between citizens and non-citizens (calculated as the quotient of the coefficient-related part of the decomposition and the wage gap, in percentages). This is the highest value, yet both Hispanics and Central Europeans are groups for which the wage gap its self is larger. For Hispanics

discrimination accounts for 75% of the wage gap and for Central Europeans the respective number is 61%. 48% of the wage gap of North Africans and Southwest Asians is due to discrimination , while for West Europeans and Sub-Saharan Africans the values are 27% and 30%, respectively.

As a robustness check for this part of the analysis we have explored the discrimination hypothesis also on groups defined not by ethnicity, but by the level of development of the country of birth. Our hypothesis was that the higher the level of development, the better the quality of education, and the better the quality of education, the better the signaling effect of education on the US labor market, and the higher the wage. Persons born in highly developed countries should thus have lower propensities to become self-employed, as they would be less discriminated. Oyelere and Belton (2012) have approached this issue and found the level of development of one's country of birth is in fact positively correlated with the odds of becoming self-employed in the US. The countries of birth rather than declared ethnicities were used in this case in order to obtain clear matches with the countries recognized by the UNDP in their Human Development Indices[9]. The Oaxaca-Blinder decomposition of wage gaps between citizens and non-citizens within four groups differing in the level of the birth countries' HDI[10] led to results which were meaningless for further analysis - only in the "medium HDI" group did significant discrimination occur. Thus, the discrimination measures obtained from comparisons between aggregate ethnic groups have been used in further analyses.

Following the approach used in estimating the relation of discrimination to the intensity of self-employment among Puerto Rican immigrants, we have run an identical probit model, in which the extent of discrimination, i.e. the value of the coefficient-related part of discrimination divided by the state average wage, is used as an explanatory variable.

Table 5.8: Probit model output, marginal effects

	dy/dx
discrimination	-0.658***
age	0.015***
gender (1=female)	-0.025***
education	incomplete - reference
primary	0.001*
secondary	0.017***
tertiary	0.007***
N	52,902
English proficiency	+
occupations	+
industries	+
specification of discrimination measures	
occupations	+
earnings	yearly

* significant at 10% level, ** significant at 5% level, *** significant at 1% level

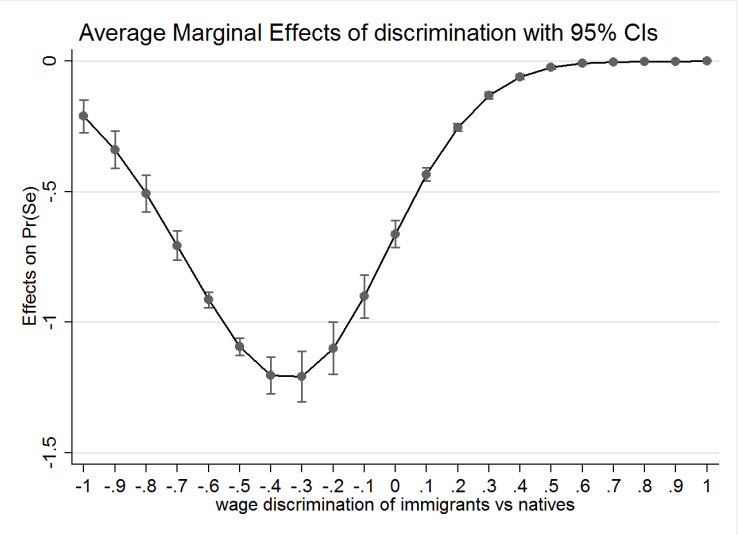
[9] http://hdr.undp.org/en/statistics/

[10] Following UNDP: very high HDI, high HDI, medium HDI, and low HDI

The estimated model indicates a significant negative effect of discrimination on the probability of becoming self-employed. This implies that as discrimination increases, the intensity of self-employment decreases. Such a result contradicts the finding for Puerto Rican immigrants. As expected, though, the explanation may be in the relatively small earnings of immigrants in self-employment.

A specific scenario, identical to the tested for Puerto Ricans, has been also tested (male, age 35, secondary education, very good command of English, working as a technician, salesman or administrative worker in the retail industry). The result was not different, i.e. the marginal effect of discrimination on the probability of becoming self-employed was -0.67 and highly statistically significant. Probability of self-employment as a function of a whole range of values of immigrant wage discrimination is presented in Figure 5.5.

Figure 5.5: Marginal effects of various levels of discrimination on the probability of immigrant self-employment in the US



5.6 Summary and conclusions

The objective of this study was to test whether the extent of labor market discrimination is significantly related to the intensity of self-employment among immigrants. First, inquiring into the migration literature we have computed a labor market discrimination measure based on an Oaxaca-Blinder decomposition of wage gaps between immigrants and natives. This measure has been computed for differences in wages between Puerto Rican immigrants in the US and four other comparison groups of natives: Puerto Ricans, Hispanics, whites, and non-Hispanic whites. On one hand, the disadvantage theory suggest that the greater the greater the social distance between two groups, the greater the extent of discrimination. On the other hand, NELM could lead us to believe that the more relevant the group of reference, the greater effect a certain comparison would have on one's labor market behavior. These considerations were critical in light of the second step of the analysis,

i.e. considering the extent of labor market discrimination as the explanatory variable in models of the intensity of self-employment. The initial hypothesis was that the higher the level of discrimination, the higher the intensity of self-employment.

The analysis has pointed to major differences in the discrimination measure when various comparison groups were considered. Nevertheless, it was possible to observe a general tendency, namely that when compared to more similar groups (Puerto Rican and Hispanic natives) Puerto Rican immigrants were rather negatively discriminated, while when compared to more distant groups (whites and non-Hispanic whites) they were rather positively discriminated (preferentially treated). Discrimination has been considered on the state level. Here some variability has also been recognized. In some states discrimination was consistently positive or negative across all comparison groups. In other states it varied, usually following the pattern related to group familiarity described above.

Based on these findings it has been concluded that institution of labor market discrimination is far more complex than expected. For one thing it varies across localities implying high context-specificity. Second, it differs greatly across comparison groups. The latter finding contradicts the intuition that the more distant the group, the greater the discrimination would be. Nevertheless, these patterns may be explained by a scenario in which the employers are in fact from the minority group, rather than from the majority group, as often assumed. This narrative calls for further scientific inquiry, especially as immigrant, and minority populations in general, grow and develop towards becoming minority-majorities.

The analysis of Puerto Rican immigrant discrimination has exploited both of the mechanisms which potentially moderate the effect of discrimination, i.e. labor market duality (by considering models which treat and do not treat occupations as a determinant of earnings) and, through the notion of multiple comparison groups, relative deprivation. When discrimination was considered as a factor related to the intensity of self-employment the findings were in line with the theory - the higher the labor market discrimination, the greater the probability that a Puerto Rican immigrant in the US will choose self-employment as an alternative labor market strategy. The most notable findings thus include the observations that:

- Puerto Rican immigrants in the US are rather positively discriminated against most distant ethnic groups (whites and Non-Hispanic whites);
- Puerto Rican immigrants in the US are rather negatively discriminated against most similar ethnic groups (Puerto Rican natives and Hispanics);
- Being discriminated against in comparison to more similar groups (native Puerto Ricans or Hispanics) has a relatively small effect on the intensity of self-employment;
- In relation to more distant groups (whites or Non-Hispanic whites) the relation of the extent of this discrimination to the intensity of Puerto Rican immigrants' self-employment across US states was relatively large.

The extension to this study aimed at enabling an extrapolation of the results obtained for Puerto Ricans to a more general setting, i.e. of immigrants in the US *en masse*. The Oaxaca-Blinder decompositions have compared the earnings of citizens and non-citizens across seven major ethnic groups, delimited more or less by continental divisions: West Europeans, North Africans and Southwest Asians, Hispanics, Central Europeans, Sub-Saharan Africans, Asians, and North Americans. Here the discrimination measure also pointed to significant differences. West Europeans were most interesting, as they were the only group within which immigrants earn significantly more than the natives and in which immigrants are positively discriminated. Due to methodological and data limitations the analysis was performed on the country level.

The effect of various groups' discrimination on the average intensity of immigrant self-employment has been found negative, contrary to the result obtained for Puerto Ricans. This result contradicts the theory based on which we should expect discrimination to push immigrants into self-employment. An explanation for this "odd" result can be found in the theory of employer discrimination, though. By means of the Propensity Score Matching technique the average earnings in wage-employment and self-employment have been compared for the immigrants, and for the natives. The obtained results lead to concluding that discrimination does not push immigrants into self-employment because the immigrants' earnings from self-employment are relatively lower than their earnings from wage-employment.

Despite the very consciously applied methodology, this study potentially suffers from a number of limitations. One of the most common weak points of studies of this type, and generally of the type presented in this thesis, are self-reported earnings. This may be a serious issue especially when comparing the wage-employed to the self-employed, as the latter may have more possibilities of affecting their earnings by, e.g. avoiding taxes or performing additional, unreported activity. Also when comparing natives to immigrants, the latter may be unwilling to declare their income from unregistered economic activity. Given the inability to correct for this potential bias, the earnings used are assumed to be proxies of actual earnings nevertheless. An other issue is that of comparison groups used in the analysis. Whether Puerto Rican immigrants actually compare their earnings to those of the suggested ethnic groups and whether immigrants and natives within the proposed ethnic clusters actually care for the differences in each others wages is unknown. The proposed empirical approach has been developed based on general theories related to inter-group relations rather than actual inquiries into how people make their decisions. Notwithstanding this approximation of reality, the obtained results are in line with intuition and allow for a rather coherent narrative. The third possible limitation arises from the applied measure of discrimination. It is as valid, as are the wage regressions. Unfortunately it is rather impossible to account for all possible productivity-related explanatory variables. Nevertheless, the fact that the dataset used is very rich most certainly improves the estimates to the greatest extent possible.

The diversity of results obtained from this study, visible on regional level as well as in relation to various ethnic groups, is surely one of the possible fields of further academic inquiry. It also touches upon a wider problem of the usefulness, on one hand, of case studies, which may be very specific and not alike any other example, and, on the other hand, the usefulness of aggregate inquiries, which clearly conceal the actual complexity of social phenomena.

Chapter 6

Competition and complementarity

6.1 Introduction

In the previous chapters we have explored how profitable is choosing self-employment as a market strategy for migrants by comparing the earnings of observably similar individuals. Given the results of that exercise we have further inquired into the institutional factors which may push individuals into self-employment. Having concluded that in the case of Puerto Rican immigrants in the US discrimination does in fact positively relate to their propensity to become self-employed, we now move to exploring yet another factor affecting the profitability of immigrant entrepreneurship, namely competition and complementarity generated by the local pool of one's co-ethnics. The objective of this chapter will be to test the following hypothesis

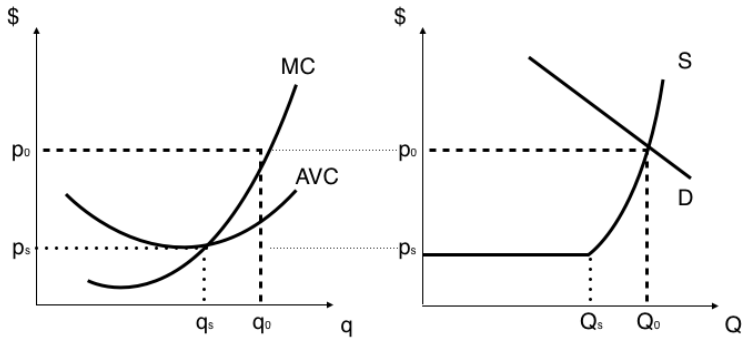
Hypothesis 6 *Ethnic competition decreases, while ethnic complementarity increases returns to local entrepreneurial activity.*

The presumption that local pools of co-ethnics may affect local ethnic business profitability arises from interrelating the postulations of two strands of literature. One is the discourse on ethnic entrepreneurship in which the main contributions are works by Wilson and Portes (1980), Wilson and Martin (1982), Sanders and Nee (1987), Portes and Jensen (1992), or Sanders and Nee (1992). These authors proposed that relatively high concentrations of co-ethnics in given localities (labeled as ethnic enclaves, niches, or economies) allow entrepreneurs to gain access to labor and financing, which are prerequisites for a successful startup. People of common cultural heritage may also be the source of increased demand for traditional products and services. Auster and Aldrich (1984) have relaxed the spatial concentration requirement. On one hand they do mention the advantages of operating a business in an enclave, where the entrepreneurs can, better than mainstream businessmen, respond to the local customers' ethnic tastes. On the other hand, though, they mention that enclaves actually do not have to denote a place, but rather "networks of communal solidarity" (Auster and Aldrich, 1984, p. 53), which can be spread across distant areas. This communal solidarity, as it is claimed, could increase survival rates of

immigrant businesses by providing support and stability in the form of preferential loans, business sites, supplies etc.

In the context of favorable, business enabling, and stimulating interconnections between enterprises, the theory of industrial organization provides counterarguments to those posed by theorists of ethnic entrepreneurship. Models of competitive markets imply that the more competition a firm faces, the smaller is its market power, i.e. the ability to set the price above marginal costs and gain substantial profits (Carlton and Perloff, 2005). Under the assumptions of perfect competition, a homogenous, perfectly divisible product, perfect information, no transaction costs, no externalities and firms being price-takers the following figures show possible profits from business operation.

Figure 6.1: Short-run competitive equilibrium: representative firm (left), market (right).



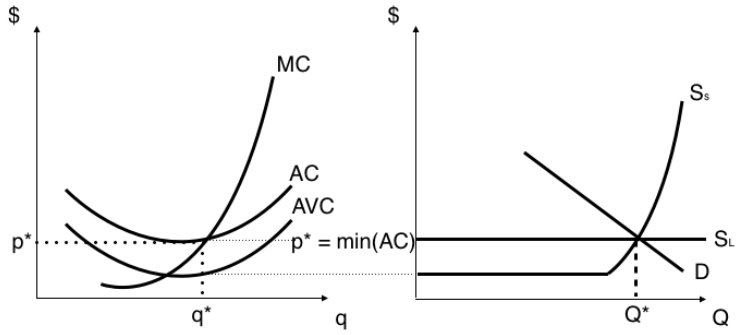
Source: Carlton and Perlof (2005), p. 62.

In the short run (Figure 6.1) all firms produce as long as the market price is above the shut down price (p_s), at which it is not profitable for any firm to generate output given the related costs. Taking the market price as given, and above p_s, firms generate profits of magnitude equal to the difference between the price and the average costs (AC). Had the price exceeded AC but not the average variable costs (AVC) firms would stay in business despite generating loss. The loss would then still be smaller than if the company was to shut down and face the loss of average fixed costs, which are assumed to be sunk. The representative firm's profits are thus equal to: $\pi = (MC(q_0) - AVC(q_0)) \times q_0$. The opportunity of making a profit creates an incentive for other companies to enter the market. This has its consequences in the long-run (Figure 6.2).

Assuming that all the incumbents and entrants are identical in terms of their marginal and average cost of production, and they employ only a small fraction of the total amount of factors of production, the long-run supply curve becomes a horizontal line at the point where the average costs are minimal, what also indicates the equilibrium price, implying zero profits. Thus the theorization that the more firms compete over the market, the smaller (infinitesimal) the profit of a typical firm.

The assumption that we are dealing with a number of identical firms is quite far-fetched of course. Let us assume that there are two types of firms which exhibit the same marginal cost functions, but different average cost functions (for simplicity let us label them 'high'

Figure 6.2: Long-run competitive equilibrium: representative firm (left), market (right).



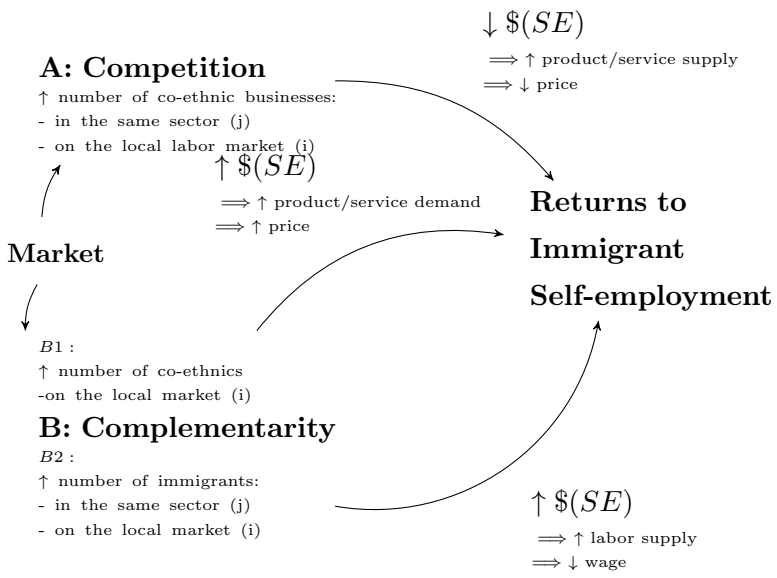
Source: Carlton and Perlof (2005), p. 62.

and 'low'). When the low cost firms reach their production capacity (they produce at minimum AC and generate no profits) and demand is not yet satisfied, producing extra units of a good implies higher average costs. The price increases to cover those costs. At that point, though, the high cost firms for which the minimum of their cost function is at the level of the new, higher market price may enter the market as production becomes feasible for them.

This logic, still being an oversimplification of reality, has lead to thinking that it is possible that due to the different cost functions and unsatisfied demand a variety of companies may actually exist on the market optimizing their profits by producing different amounts of output. In such a scenario the low-cost companies should be able to generate extra rent, as compared to their high-cost counterparts. Thus the more unequal the distribution of market shares, the higher the profits can relatively be (Tirole, 1988). Research in the field of industrial organization has found empirical evidence of this relationship. It has been shown that indeed higher market concentration (more unequal distribution of market shares, the extreme of which is a monopoly facing no competition) leads to higher industry profitability (Bain, 1956; Demsetz, 1973; Cowling and Waterson, 1976; Tirole, 1988). The link holds also when the market power of industries is accounted for (hypothesizing the possibility of charging higher prices in more concentrated settings, Peltzman, 1977). Whether the concentration-profitability principle applies also to ethnic business operation, as described by migration theorists, and how ethnic businesses fare among within-industry competitors are inquiries not yet tackled, though.

By adding the *ethnic* to the *business* we will be able to capture, otherwise unrecognized, advantages which arise from the complementarity of business owners and their clientele and workers. Recognizing the preferences of ethnic clients and enabling relational contracting with ethnic employees may significantly decrease the costs of business operation and increase the profits. The greater the number of firms which recognize this source profitability, though, the grater the competition and the smaller the profits. A theoretical framework encompassing the effects described above is presented in Figure 6.3.

Figure 6.3: Effects of market competition and complementarity on immigrant self-employment.



In applying the concepts of market competition/concentration to ethnic entrepreneurship we will be continually considering the case of Puerto Ricans in the US. This leads to a definitional issue, which was not as critical in the previous two sections. Hitherto we have analyzed either individual features of immigrants and how they relate to individual outcomes (in the study on income gains to self-employment), or, when considering a contextual factor such as discrimination, we did not require that ethnic origin play any other role than that which we actually capture, i.e. that an individual was born in Puerto Rico and considers himself or herself of Puerto Rican decent. In this chapter we have to make stronger assumption concerning the role of ethnicity in business operation. Based on three facts - that an individual is of Puerto Rican ancestry, was born in Puerto Rico and is self-employed in the US in a given industry - we take it for granted that an individual would compete with his or her co-ethnics, as if they operated in an ethnic economy. This of course must not be true, as even a Puerto Rican operating a grocery store must not necessarily supply Puerto Rican food and compete with other Puerto Rican owners of "Puerto Rican" stores. Instead it is possible that they run a store with Indian food and supply local Indian (or any other as a matter of fact) customers and compete with Indian (or any other) owners of similar stores. Waldinger et al. (1990) in their seminal work on ethnic entrepreneurs state that

> what is "ethnic" about "ethnic business" may be no more than a set of connections and regular patterns of interaction among people sharing common national background or migration experiences (Waldinger et al., 1990, p.33).

Though we do actually consider people "sharing common national background or migration experiences", the first part of the above-mentioned quote is what we assume to

be the case, rather than what we can actually observe. Unfortunately, the data does not allow for the recognition of such interactions. Instead we rationalize this assumption by claiming that if a Puerto Rican runs a business (what we observe) which is a grocery store with Indian food products, serving the Indian community in an Indian neighborhood (what could be the actual story behind our data) he or she could just as well run a store oriented towards Puerto Ricans, as this is what is his or her inherent and inalienable advantage. Nevertheless, all conclusions derived from this chapter should be interpreted with this inference in mind.

6.2 Methodology

The methods used to answer the questions posed above include computing the Hirschman-Herfindahl Index (HHI) of industrial concentration, and estimating Mincerian wage equations, which account for the non-random selection of Puerto Ricans into self-employment, and using the HHI as a contextual explanatory variable. Both procedures are described in detail in the following sections. A description of data transformations conducted for the purpose of this study follows.

6.2.1 Measures of industry concentration

Measures of industry concentration are measures used to determine the distribution of the market shares of specific firms. These measures can be most basically divided into those that account for all companies on the market, and into those which consider only few (Ginevicius and Cirba, 2007). Their overview is presented in Table 6.1.

Table 6.1: Market concentration measures

Measure	Functional form
Concentration Ratio	$\mathrm{CR}m = \sum_{i=1}^{m} s_i$
Hirschman-Herfindahl Index	$\mathrm{HHI}=\sum_{i=1}^{n} {s_i}^2$
Horwath Index	$\mathrm{HI}= \mathrm{s}_1 + \sum_{i=2}^{n} {s_i}^2(2 - s_i)$
Entropy Index	$\mathrm{EI}=-\sum_{i=1}^{n} s_i ln s_i$
Exponential Concentration Index	$\mathrm{ECI}=\Pi_{i=1}^{n} s_i^{s_i}$
Rosenbluth Index	$\mathrm{RI}=1/[(2\sum_{i=1}^{n} i s_i) - 1]$
GIN Index	$\mathrm{GI}=\sum_{s=1}^{n} s_i/[1 + n(1 - s_i)]$

m - number of largest firms, i - index of i-th largest firm, s_i - market share of firm i, n - total number of firms,

Among measures which capture only part of the distribution of market shares we can find the concentration ratios (CRs), which indicate the share of industry market sales of, most commonly, 4 (CR4) or 8 (CR8) largest firms. The ratios fall in the range $(0\%, 100\%]$, where 100% means the given number of largest companies are the only suppliers on the market. This of course indicates the biggest drawback of concentration ratios, namely

that they are not applicable if the number of companies is smaller than n, and that they are informative as far as the competition within the top n firms is concerned.

The other measures summarized in Table 6.1 are functions of market shares of all the companies within certain industries. This is preferential in comparison to the CRs, but as can be seen the weights attached to particular sizes of firms significantly vary and affect the value of the indices. One of the most frequently used alternative measures is the Hirschman-Herfindahl Index (HHI). The square of the companies' shares results in grater weights given to larger firms. Thus, it can capture not only the dispersion related to the number of firms, but also the inequality of their market shares. HHI makes it possible to distinguish between three companies claiming 50%, 25% and 25% of the total sales, and three companies which all have equal market shares ($\sim$33%). In the first case the larger company (50% share) receives a higher weight than the other two, thus the value of the index increases relatively to the case where the three companies have identical sales. The index takes values in the range $(0, 1]$, where 1 indicates that a company is the only firm in a given industry, a monopolist. The notion that the HHI uses square terms, i.e. weights which are proportional to the market share of a company, has been contested and thus the other measures have been proposed.

The Horwath Index (HI) considers the share of the largest company as it is and for the smaller firms it weighs them proportionally to their size (square term), and additionally applies a weight which inversely proportional to the size $(2 - s_i)$. These operations result in the fact that, contrary to HHI, HI tends to take values in the middle and top of the value range (0,1]. Why the largest firm has been singled out (rather than e.g. two or three largest firms) seems quite arbitrary, though. As a matter of fact the division into the discrete and cumulative parts is itself controversial (Curry and George, 1983; Ginevicius and Cirba, 2007).

Two alternative measures of industrial concentration which, similarly to the HHI, are more concise in their form are the Entropy Index, and the Exponential Concentration Index. The former relates the concept of (information) uncertainty to market outcomes. It assumes that as competition grows, the probability that a customer will provide business to a given company decreases. A monopoly would have lowest entropy, as the firm is perfectly informed, i.e. able to predict customer's actual choices, and does not face any uncertainty. The intuition laying behind applying this measure seems quite controversial, however, as it originates in a very different field. Its properties also imply that, as opposed to all the other concentration measures, it decreases as the market structure approaches a monopoly. The ECI has been found quite similar to the HHI (Ginevicius and Cirba, 2007), the only difference being the weight assigned to small and large firms. While the HHI was found by Ginevicius and Cirba more sensitive to large firms, the ECI is more sensitive to small firms. Thus, as the authors claim, in order to estimate the ECI the whole distribution of market shares has to be known, while in order to obtain a decent approximation of HHI knowing the low end of the share distribution is not critical.

The final two measures considered in Table 6.1 are the Rosenbluth Index, and the GIN Index. Due to the product of if the i-th firm's share and i itself in the denominator of the former, it gives different weights to companies which have identical shares. This, even intuitively, seems like a disadvantageous property. The latter measure captures both the number of firms (n), and the relative size of every firm (s_i) in evaluating their relative sizes (used in the summation). As Ginevicius and Cirba (2009) claim in a later text of theirs, it underreports competition in case of e.g. a market comprising two companies, one of which has 90% of the market shares, and the other one 10%. The value of the GIN index is then 0.786 - much smaller than an expected 0.9. The HHI in this case would give the value of 0.82.

What stems from the above discussion is a conclusion, that neither measure of industrial concentration can be considered to perform best based on objective criteria. The discourse does not provide a recommendation for which index to use. While some of them are more suitable for very dispersed markets, others better capture competition among few large firms. Given the incomplete coverage of the concentration ratios we have rejected their applicability in the first place. The debatable features of other measures be it their theoretical grounding (in case of the Entropy Index), a large degree of arbitrariness in choosing the functional form (Horwath Index, Rosenbluth Index), or strong assumptions concerning complete market information (Exponential Concentration Index) have discouraged us from employing them in this study. Out of the reviewed measure only the HHI and the GIN Index have been eventually considered. Since they perform in a relatively similar manner, both nevertheless facing some drawbacks, following Ockham's Razor, HHI as the simpler (and more commonly used) one has been eventually applied.

Bearing in mind the disadvantages resulting from the functional form of the HHI, another debatable aspect of this (and any other concentration measure) should be noted, namely the attribute of a company which is to be used to compute its market share. Carlton and Perloff (2005) consider the value of sales as a default. Ginevicius and Cirba (2007) refer to "attributes" in general, without providing tests of the sensitivity of reviewed measures with respect to different characteristics. Curry and George (1983) mention that among the proxies for a firm's size we could consider value-added, sales, employment or assets. The choice for this study was limited by data availability. As it will be described in the data section, one's self-employment, pre-tax income has been used as the company performance measure.

6.2.2 Mincerian wage equations

A standard approach in the literature on the returns to self-employment (Li, 2000; Lofstrom, 2002; Portes and Shafer, 2006) consists of estimating a Mincerian-type equation

$$\text{Income of self-employed}_{i,r,k} = \beta_0 + \beta\mathbf{X} + \beta_r\mathbf{r} + \beta_k\mathbf{k} + \epsilon_i, \tag{6.1}$$

where income is measured on comparable basis (e.g. per hour or per year) across individuals i, sectors k, and across r localities. Matrix $\mathbf{X}$ typically comprises standard explanatory

variables such as age, gender or education. The identification strategy proposed in this study extends the standard specification by additional control variables, which capture the effects of general competition, ethnic competition, and ethnic complementarity across markets (sectors and geographic localities). These include:

- **industrial competition** measured by the Hirschman-Herfindahl Index across industries k and local areas r.;
- **ethnic competition** measured by the Hirschman-Herfindahl Index encompassing only Puerto Rican businesses within each sector k and locality r;
- **ethnic complementarity** measured by the density of Puerto Ricans relative to non-Puerto Ricans, as well as the relative wealth of Puerto Ricans relative to non-Puerto Ricans in area r.

Technical reasons which prevent the use of OLS in estimating returns to self-employment are twofold. First, there may be systematic *selection into* self-employment. Second, there may be also systematic success factors, whereas typically large (positive and negative) incomes of the self-employed are censored. These two problems necessitate the use of the correction for usually latent components of returns to self-employment (Simpson and Sproule, 1998; Le, 1999; Åstebro and Chen, 2012).

Following (Berglann et al., 2011; Hamilton, 2000) a two-stage procedure is formulated. We use

$$Pr(working) = \Gamma_0 + \mathbf{\Gamma X} + \varepsilon \tag{6.2}$$

to obtain estimates of the Inverse Mill's Ratio (IMR) in order to correct the bias in the estimates of (6.1). The variable used to identify the selection equation, identically to the approach applied in the previous chapter, is one's marital status (see (Cameron and Trivedi, 2009) for a discussion of using or not using an identification variable). We then move to estimating

$$\begin{aligned} SE\ income \quad = \quad & \beta_0 + \beta \mathbf{X} + \beta_r \mathbf{r} + \beta_k \mathbf{k} + \delta IMR + \\ & \gamma_1 General\ Competition + \gamma_2 Ethnic\ Competition + \\ & \eta Complementarity + \epsilon_{Income}, \end{aligned} \tag{6.3}$$

for all positive incomes from self-employment[1].

6.2.3 Data

The models for this section have been estimated on the same data as the estimations in the previous chapter, the 5-year 1% Public Use Micro Samples of the American Community Survey, what is a dataset encompassing ACS data from the years 2005-2010 (Ruggles et al., 2010). Additional data transformations include computing the HHI on different levels of

[1] We do not take into account negative values of income. The primary reason for doing so relies in how HHI is computed - negative values would lower the overall shares of income.

geographic aggregation, and computing ratios of co-ethnics and of co-ethnic wealth on different levels of geographic aggregation.

The HHI has been computed for 113 industries[2]. The ethnic, Puerto Rican HHI (HHI_{PR}) is computed for persons of Puerto Rican ancestry operating businesses in those 113 industries. Controls for either 533 Super Public Use Microdata Areas[3], or, in the alternative specification, 52 states[4] have been considered.

In order to capture the complementarity effects depicted in Figure 6.3, ratios of persons of Puerto Rican origin to the total population of a given Super-PUMA or state have been calculated and used as contextual, explanatory variables in the models. The relative wealth, approximated by average total household income, of the local Puerto Rican community in relation to the average wealth of the total population has also been accounted for in the models. As far as immigrant specificity is concerned, an *immigrant* dummy variable equal to 0 if one is of Puerto Rican ancestry[5] and US born, and 1 if one is of Puerto Rican ancestry and was born in Puerto Rico[6].

Employed measures of income are identical to those used in previous chapters. Basic descriptive statistics for the total sample and two Puerto Rican subgroups - persons of Puerto Rican ancestry born and residing in the US, and persons of Puerto Rican ancestry born in Puerto Rico and residing in the US are presented in Table 6.2. The last column of Table 6.2 presents the differences in explanatory variables between Puerto Rican "natives" (2) and Puerto Rican "immigrants" (3)[7].

[2]Three-digit level of the classification yielding 196 categories - the *IND1990* variable in the USA-PUMS database. All estimations have been run without supplied weights - multiplying observations lead to possibly "fake" concentration indices. The unweighted sample did not present significant distributional differences. We have also excluded the Agriculture, Forestry, Fishing and Hunting and the Mining industry, as well as the Public Administration and the Military as there were negligible numbers of self-employed in those sectors, what eventually biased the concentration index estimate.

[3]Geographic areas with min. 400,000 residents - the *PUMASUPR* variable in the USA-PUMS database.

[4]The *STATEFIP* variable in the USA-PUMS database.

[5]Based on the first response to the question about one's self-reported ancestry or ethnic origin, the $ANCESTR1$ variable in the PUMS database.

[6]Differentiation made based on the declared birth place of every respondent, the BPL variable in the PUMS databases.

[7]Recall that Puerto Ricans are technically US citizens and thus they are not recognized as immigrants in the American Community Survey questionnaire.

Table 6.2: Summary statistics - means or proportions

	(1) All	(2) US born Puerto Ricans	(3) Puerto Rico born Puerto Ricans	Difference (2)-(3)
Income per hour	18.088	18.572	17.401	
	(22.308)	(22.774)	(21.620)	
age	43.575	40.431	47.667	**
16-75	(12.743)	(11.639)	(12.957)	
men	58%	58%	58%	
women	42%	42%	42%	
education-incomplete	1%	1%	2%	**
education-primary	5%	2%	9%	**
education-secondary	48%	49%	46%	
education-tertiary	46%	48%	43%	**
English-does not speak	2%	0%	3%	**
English-speaks not well	6%	1%	11%	**
English-speaks well	12%	6%	21%	**
English-speaks very well	52%	50%	55%	**
English-speaks only English	28%	43%	10%	**
HHI by Super-PUMA	0.125	0.125	0.124	
	(0.185)	(0.183)	(0.187)	
HHI by state	0.013	0.014	0.012	
	(0.044)	(0.049)	(0.035)	
HHI_{PR} by Super-PUMA	0.770	0.383	0.738	**
	(0.306)	(0.357)	(0.322)	
HHI_{PR} by state	0.362	0.794	0.334	**
	(0.3512)	(0.292)	(0.342)	
Relative number of co-ethnics	0.043	0.039	0.048	**
by Super-PUMA	(0.057)	(0.055)	(0.058)	
Relative number of co-ethnics	0.022	0.021	0.023	**
by state	(0.014)	(0.014)	(0.013)	
Relative wealth of co-ethnics	0.818	0.835	0.796	**
by Super-PUMA	(0.212)	(0.212)	(0.210)	
Relative wealth of co-ethnics	0.777	0.787	0.766	**
by state	(0.178)	(0.173)	(0.183)	
non Puerto Rico born	57%	n/a	n/a	
Puerto Rico born	43%	n/a	n/a	
No. of observations	3298	1865	1433	

Data: IPUMS database - ACS 5 year sample 2005-2010. Standard deviations in parenthesis. (**) implies difference in means or difference in percentage points, where applicable, significant at 5% level.

For the purpose of modeling and meaningful interpretation of the results, all variables (but dummies) have been standardized. Otherwise a unit increase in the competition measures would have been incomprehensible.

6.3 Self-employment returns and ethnic congestion

Obtained results are OLS estimates and are corrected for non-random selection of individuals into working by means of the Heckman correction. To facilitate interpretation, all continuous variables have been standardized. Consequently, point estimators may be interpreted in terms of relative importance.

Table 6.3: Regression output. Super-PUMA level of analysis.

	(1)	(2)	(3)	(4)
Competition effects:				
HHI	-0.030	-0.030	-0.029	-0.067**
$HHI_{\text{Puerto Rican}}$				0.073***
Complementarity effects:				
relative number of co-ethnics		-0.077***	-0.066***	-0.032
relative wealth of co-ethnics			0.041**	0.036*
Standard covariates:				
age	0.978***	1.313***	1.346***	1.368***
age^2	-0.666**	-0.970***	-1.003***	-1.021***
gender (1-female)	-0.147	-0.115	-0.111	-0.128
age **x** gender	-0.072	-0.095*	-0.097*	-0.092
education - incomplete		reference level		
education - primary	-0.007	-0.020	-0.027	-0.010
education - secondary	0.075	0.156	0.155	0.152
education - tertiary	0.483***	0.574***	0.569***	0.550***
English - does not speak		reference level		
English - speaks not well	-0.156	-0.132	-0.132	-0.119
English - speaks well	-0.168	-0.149	-0.142	-0.138
English - speaks very well	0.047	0.073	0.075	0.076
English - speaks only English	0.072	0.071	0.068	0.063
Inv. Mill's Ratio	0.022	0.192	0.213	0.206
constant	-0.540	-0.792**	-0.813**	-0.816**
No. of observations	2853	2853	2853	2853
R^2	0.092	0.096	0.097	0.101

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

We have been concerned that the results were driven by the Super-PUMA level of analysis. In order to explore this issue we have run an identical set of models, only that all spatial measures were calculated on the state level. The results did not present statistically significant effects in terms of the variables of primary interest. This most probably resulted from insufficient variation on such a high level of aggregation.

Results obtained for Super-PUMAs present several interesting patterns. Firstly, it can be observed the estimator of the Inverse Mill's Ratio is insignificant, implying no relevant selection bias of Puerto Ricans into working. This result is somewhat puzzling given results obtained from models with identical selection restrictions. Selection equations estimated with marital status as a selection variable have given ultimately significant Inverse Mill's

Ratios in other studies of self-employment (see e.g. Tyrowicz and Nestorowicz, 2011), including the extension to this analysis presented in the latter chapter of this thesis. Secondly, thanks to modeling standardized variables we can see that *age* is the variable which explains most in terms of the variation in returns to self-employment in the analyzed sample. A one standard deviation increase in age is related to approximately one standard deviation increase in returns to self-employment. As age^2 is also significant, and negative across all models we may conclude that there exists a non-linear relationship between age and returns to self-employment such that they are increasing with a decreasing rate. All but one of the other control variables are insignificant. It appears that in the case of entrepreneurs of Puerto Rican ancestry working in the US not gender, nor proficiency in English are significantly related to returns to self-employment. Only having tertiary education, as compared to not completing even the level of schooling, is associated with higher earnings from entrepreneurial activity.

The estimators of greatest interest are those measuring the competition and complementarity effects. Overall competition within industries showed no significant relation to returns to business operation for Puerto Ricans. This may be on hand due to the relatively small number of Puerto Rican entrepreneurs, and on the other hand to the incubating effect of ethnic enclaves which, according to theorists of ethnic entrepreneurship, may protect the ethnic economy from the outer market. HHI becomes significant, though, in the last model, where we also account for Puerto Rican HHI, that is within sectoral competition, but only among Puerto Rican businesses. It shows that as the general level of competition in an industry increases (HHI falls), returns to business operation increase for the Puerto Ricans, under the condition that competition among them remains the same. This result is indicative of an ethnic niche or some kind of enclave economy being in place. Businesses run by Puerto Ricans are on one hand isolated (independent) from the industrial environment in the local area (Models (1)-(3)), but once competition among Puerto Ricans themselves is accounted for (Model (4)), the local industrial relations have a beneficial effect on their profits - the more competition there is outside of the, hypothesized, "enclave", the more money the "enclave" makes. A 1 s.d. increase in general competition (fall of HHI by 0.18, equivalent to a change of the market structure from 3 equally sized companies to 10 equally sized companies) is related to a USD 1.5 per hour ($0.067 * 22.03 = 1.48$) increase in returns to self-employment among the Puerto Ricans ($\sim$USD 2,800 per year).

As far as the Puerto Rican HHI is concerned, we find that within-group business competition, ceteris paribus, i.e. assuming that industrial competition in general does not change, is detrimental to profits, as expected from the theory (the higher the HHI, the closer the market is to a monopoly, thus the higher should be the expected profits). An increase in the Index by one standard deviation (0.3) is related to an increase in profits by 0.073 standard deviations ($0.073 \times 22.03 = 1.6$). Such changes account e.g. for the following situation: if initially we have two equally sized companies (each with a 50% share of the market) and one company expands to cover 90% of the market, and the other

one is left with 10% (change of HHI by 1 s.d., 0.3, from HHI=0.5 to HHI=0.82), then the average return to business operation will increase by approximately by USD 1.6 per hour (USD 3,300 per year). Expectedly, the influence of the structure of the ethnic businesses has a stronger relation to business income then the overall structure of businesses in an industry.

When it comes to variables capturing circumstances which are presumably complementary to business operation, i.e. the number of co-ethnics in a given locality, and their wealth, their estimators turn out significant in all of the considered models. Based on the literature reviewed for the purpose of this thesis, we expected that the number of co-ethnics will be positively related to returns to one's business operation. The data indicates the contrary. One possible reason could be in the somewhat expected effect that the more co-ethnics there are, the more of them may also take advantage of their ancestry and start their own enterprise. The other explanation could lie in the fact that the larger the ethnic concentrations in a given area, the more likely it is we are dealing with ethnic ghettoization, possibly also related to relatively low economic status of the residents, what in turn results in low purchasing power. Thus in Model (3) not only the proportion of co-ethnics in the local population was accounted for, but also their relative wealth. It turns out that the richer the local pool of co-ethnics, the more income co-ethnic entrepreneurs generate. The relation is weaker, though, than that between profits and the other complementarity and competition effects. In Model (4) both competition and complementarity conditions are accounted for. Adding ethnic HHI to the model turned the estimator of the number of co-ethnics insignificant. To some extent this strengthens our supposition, that more co-ethnics may in fact relate to more co-ethnic competition in business. The relation of business returns to co-ethnic wealth remains similar in terms of both directionality and scale.

As the obtained results may be sensitive due to definitional, conceptual and methodological assumptions, two robustness checks have been carried out on the final model. Firstly, we were interested whether "ethnic monopolists", i.e. persons who are the only local Puerto Rican entrepreneur operating business in a given industry, significantly alter the results. Secondly, the notion of being an immigrant, central to this thesis, and its relation to business returns was examined. The results are presented in Table 6.4. Model (4a) excludes observations, which were characterized by either $HHI = 1$ or $HHI_{ethnic} = 1$. In Model (4b) we include a dummy variable indicating whether one is a Puerto Rican immigrant or not. As indicated in the previous studies, Puerto Ricans are US citizens. Thus the *Immigrant* dummy took the value of 1 if one was of Puerto Rican ancestry, but born in Puerto Rico (recall that the American Community Survey data are employed, what implies that the respondents are in the US at the time of the survey), and the value of 0 if one was of Puerto Rican ancestry and born in the US[8].

Model (4a) reports the results of the first robustness check, namely limiting the sample to observations for which the market concentration measures are not equal to

[8]Based on the BPL and $ANCESTR1$ variables in the IPUMS databases.

Table 6.4: Robustness checks. Super-PUMA level of analysis

	(4)	(4a)	(4b)
Competition effects:			
HHI	-0.067**	-0.106	-0.068**
$HHI_{\text{Puerto Rican}}$	0.073***	0.124***	0.061***
Complementarity effects:			
relative number of co-ethnics	-0.032	-0.015	-0.031
relative wealth of co-ethnics	0.036*	-0.013	0.036*
Robustness check variables:			
HHI< 1 and $HHI_{\text{Puerto Rican}} < 1$	no	yes	no
immigrant (0-no, 1-yes)			-0.018
immigrant **x** $HHI_{\text{Puerto Rican}}$			0.027
Standard Covariates:			
age	1.368***	0.876*	1.358***
age^2	-1.021***	-0.618	-1.009***
gender (0-male, 1-female)	-0.128	-0.168	-0.131
age **x** gender	-0.092	-0.069	-0.090
education - incomplete		reference level	
education - primary	-0.010	-0.210	-0.001
education - secondary	0.152	0.017	0.152
education - tertiary	0.550***	0.400**	0.549***
English - does not speak		reference level	
English - speaks not well	-0.119	0.133	-0.122
English - speaks well	-0.138	-0.003	-0.148
English - speaks very well	0.076	0.148	0.061
English - speaks only English	0.063	0.151	0.045
Inv. Mill's Ratio	0.206	0.100	0.202
constant	-0.816**	-0.554	-0.789**
No. of observations	2853	1211	2853
R^2	0.101	0.110	0.101

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

1 (monopoly). This cuts the number of observations by more than 50%. This is not surprising due to the fact that in the base-case models (1-4) we measure market concentration within industries and Super-PUMAs, what given the initial sample size of nearly 3,000 observations must result in many cases where someone is the only Puerto Rican in a given industry and locality (please recall, though, that observations are not weighed for methodological reasons). Compared to Model (4), in Model (4a) we can not conclude that the estimator of HHI is significantly different from 0. Competition among Puerto Ricans remains significant, though, and the correlation is nearly twice as strong - an increase in $HHI_{\text{Puerto Rican}}$ by 0.3 (recall this implies moving towards a monopoly) is related to a a 2.75 USD increase in hourly earnings (5,600 USD per year). This relation is still weaker, though, than that of income and age, or tertiary education.

Model (4b) includes a dummy variable capturing whether one came to the US from Puerto Rico during his or her lifetime, or was born there. This robustness check does not meaningfully alter any of the estimators obtained in Model (4). The *immigrant* dummy variable, and its interaction with Puerto Rican HHI are not statistically significant, implying that 'immigrant' status is not relevant for the size of business income of Puerto Ricans living in the United States.

6.4 External validity of the study

In order not to limit the analysis to Puerto Rican immigrants only, the analysis of ethnic competition and complementarity and their effect on returns to self-employment are also considered across other ethnic groups, similarly to the study of labor market discrimination

The data used for this analysis is from the the American Community Survey for 2010 (1% sample of the population) and the 5% sample of the Census for 2000. The reason why these samples were used rather than the 5-year sample used heretofore is purely organizational and not expected to affect the results significantly. Just as the 5-year samples, the use of the 2010 sample, as the most recent one, and the 2000 sample as one from before the recent financial crisis assure that the results are not distinctively driven by crisis-related anomalies. In the dataset used for this study we consider 113 industries[9], controls for 533 Super Public Use Microdata Areas[10], and 8 ethnicities. Categories for ethnicity recognize the ethnic groups clustered into approximately continental divisions[11]. The measure of income is identical to the one used in the Puerto Rican study, i.e. self-reported earnings adjusted for the number of hours worked[12]. Incomes are inflation-adjusted to 2010 US dollars. The sample is restricted to adult and economically active individuals aged 16-75, but the the ratio of co-ethnics and their wealth are calculated for the whole population. Table 6.5 presents the basic descriptive statistics of the variables used in the analysis.

Similarly to the Puerto Rican study, in order to facilitate interpretation, all continuous variables have been standardized. Consequently, point estimators may be interpreted in terms of relative importance.

In almost all specifications more concentration is associated with higher profits, which is a standard finding, expected based on theory. However, ethnic HHI in Model 5 turns the overall HHI insignificant and has a negative and statistically significant sign. A large number of co-ethnics is detrimental to profitability (Model 3), but if co-ethnics are relatively wealthy, there are seizable positive effects on self-employed revenues. The results are robust across the specifications.

[9]Three-digit level of the classification yielding 196 categories - the IND1990 variable in the USA-PUMS database. All estimations have been run without supplied weights - multiplying observations lead to possibly fake concentration indices. The unweighted sample did not present significant distributional differences. We have also excluded the Agriculture, Forestry, Fishing and Hunting and the Mining industry, as well as the Public Administration and the Military as there were negligible numbers of self-employed in those sectors, what eventually biased the concentration index estimate

[10]Geographic areas with min. 400,000 residents - the PUMASUPR variable in the USA-PUMS database

[11]Considered ethnic groups include: Others, West Europeans, North Africans and Southwest Asians, Hispanics, Central Europeans, Sub-Saharan Africans, Asians, and North Americans - based on the ANCESTR1G variable in the USA-PUMS database.

[12]Reported values are censored. In the 2000 sample incomes at or above 126,000 USD are expressed as the state means of values above 126,000 USD. In the 2010 sample censuring is made at the 99.5th percentile within each state (higher values are the state means of all cases above these cutoff values). The lower boundary is imposed on 0 USD.

Table 6.5: Summary statistics - means or proportions

	(1) All	(2) Natives	(3) Immigrants	(4) Difference (2) - (3)
Income per hour	22.86	23.26	17.04	6.22**
	(29.03)	(29.38)	(22.65)	
Age	46.35	46.71	41.16	5.55**
(only 16+75 population)	(12.38)	(12.37)	(11.24)	
Gender	0.36	0.36	0.39	0.03**
(1-female)	(0.48)	(0.48)	(.49)	
Marital status	0.72	0.72	0.68	0.04**
(1-married)	(0.45)	(0.45)	(0.47)	
Education-primary	3.14%	2.25%	15.88%	-13.63 pp**
Education-secondary	41.22%	41.17%	37.73%	3.44 pp**
Education-tertiary	55.63%	56.19%	41.69%	14.50 pp**
English-does not speak	0.68%	0.10%	9.15%	-9.04 pp**
English-speaks only English	83.71%	88.35%	15.99%	72.36 pp**
HHI	0.0013	0.0013	0.0013	
	(0.0066)	(0.0066)	(0.0073)	
HHI_{ethnic}	.0008	.0005	.0045	-0.040**
	(0.0115)	(0.0065)	(0.0374)	
Relative number of co-ethnics	0.007	0.006	0.019	-0.012**
	(0.0076)	(0.0064)	(0.013)	
Relative wealth of co-ethnics	1.09	1.10	0.99	0.11**
	(0.22)	(0.21)	(.24)	
Dominance	0.14	0.13	0.24	-0.11**
(0-minor group, 1-major group)	(0.34)	(0.34)	(0.4)	
No. of observations	755,888	707,484	48,404	

Data: IPUMS database - ACS 2010, US Census 2000. Standard deviations in parenthesis.
(**) implies difference in means or difference in percentage points, where applicable, significant at 5% level.

A one standard deviation increase in the Hirschman-Herfindahl Index is related to a relatively small (0.009 s.d.) increase in returns to self-employment. The effect of competition is smaller than any of the individual level determinants included in the specification. However, HHI is much less dispersed than incomes, what implies that an increase in the value of the HHI by ten standard deviations (0.06), what corresponds e.g. to the market being shared equally between 12 rather then 40 identical firms, is related to an increase in the owners' weekly earnings by \$2.6 ($\beta_{HHI} \times s.d._y \times 10 = 2.6$). Comparatively, having tertiary rather than primary education relates to returns which are higher by \$13.5.

On the other hand a one standard deviation increase in co-ethnic concentration is associated with a 0.011 s.d. decrease in income from self-employment. Potentially, the relatively bigger ratio of co-ethnics to non co-ethnics could be associated with smaller purchasing power of members of ethnic groups clustered in enclaves. If combined with the wealth of the co-ethnics, the coefficient on the size of ethnic group turns insignificant. We can thus say that it is not the number of co-ethnics *per se*, what generates a complementarity effect with respect to entrepreneurial returns, but rather the purchasing power of one's own ethnic group.

A significant coefficient on ethnic HHI turns the estimator of HHI insignificant. Apparently, an increase in ethnic concentration fosters returns to self-employment. Entrepreneurs representing the same ethnic group, who operate businesses in a given in-

Table 6.6: Regression output.

	(1)	(2)	(3)	(4)	(5)
Competition effects:					
HHI	0.009***	0.008***	0.009***	0.008***	-0.002
HHI_{ethnic}					-0.013***
Complementarity effects:					
relative number of co-ethnics			-0.011***	0.000	
relative wealth of co-ethnics				0.027***	
Standard covariates:					
Age	1.349***	1.342***	1.346***	1.341***	1.356***
Age^2	-1.257***	-1.254***	-1.255***	-1.253***	-1.265***
Gender (1-female)	-0.118***	-0.146***	-0.120***	-0.120***	-0.121***
Age **x** gender	-0.149***	-0.149***	-0.149***	-0.148***	-0.149***
Education - primary			reference level		
Education - secondary	0.118***	0.105***	0.113***	0.108***	0.118***
Education - tertiary	0.465***	0.445***	0.460***	0.448***	0.466***
English - does not speak			reference level		
English - speaks not well	0.035***	0.024**	0.029***	0.026***	0.035***
English - speaks well	0.097***	0.079***	0.088***	0.085***	0.096***
English - speaks very well	0.286***	0.273***	0.274***	0.274***	0.286***
English - speaks only English	0.327***	0.315***	0.306***	0.311***	0.327***
Inv. Mill's Ratio	0.461***	0.462***	0.559***	0.461***	0.464***
Constant	-0.909***	-0.872***	-0.887***	-0.872***	-0.911***
Ethnic group dummies	no	yes	no	no	no
Super-PUMA dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
No. of observations	750,795	750,795	750,795	750,795	734,655
R^2	0.11	0.11	0.11	0.11	0.11

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

dustry attract clientele and may also create a support system based on ancestral linkages, e.g. enables access to informal credit lines or exchange of best business practices (e.g. Marshall-Arrow-Romer type of knowledge spillovers). Such areas as 'China Towns' or 'Little Italys', despite comprising multiple, very similar businesses enhance profitability.

The results may be susceptible to few dominant ethnicities (e.g. operating in specific industries) or some other special cases. Thus, we tried a number of robustness checks to test the vulnerability of the main estimators to the inclusion of additional controls. These include domination, knowledge of the market and separating ethnicity from immigration. Model 5a excluded those individuals who were ethnic and industrial 'monopolists' ($HHI = 1$ and $HHI_{ethnic} = 1$). In Model 5b we include a dummy variable indicating whether one's duration of stay in the US is above ($duration = 1$) or below ($duration = 0$) the median (47 years)[13]. Model 5c controls whether one's ethnic group is dominant (largest) within an industry. Model 5d considers differences in returns between immigrants and natives. Table 6.7 presents the results.

[13]The duration of stay for non-immigrants was equal to their age, for immigrants it was the length of their stay in the US as reported in the $YRSINUS$ variable in the PUMS database.

Table 6.7: Robustness checks.

	Model 5	Model 5a	Model 5b	Model 5c	Model 5d
Competition effects:					
HHI	-0.002	-0.004	-0.003	-0.005	0.006
HHI_{ethnic}	-0.017***	-0.017***	-0.009***	-0.016***	-0.050***
Robustness check variables:					
HHI< 1 and $HHI_{ethnic} < 1$	no	yes	no	no	no
Duration above median			0.025***		
Duration **x** HHI_{ethnic}			-0.030***		
Dominant (0-no, 1-yes)				0.024***	
Dominant **x** HHI_{ethnic}				0.013**	
Immigrant (0-no, 1-yes)					-0.091***
Immigrant **x** HHI_{ethnic}					0.057***
Standard Covariates:					
Age	1.350***	1.356***	1.371***	1.355***	1.347***
Age^2	-1.265***	-1.265***	-1.298***	-1.264***	-1.258***
Gender (0-male, 1-female)	-0.121***	-0.121***	-0.118***	-0.122***	-0.122***
Age **x** gender	-0.149***	-0.149***	-0.153***	-0.149***	-0.148***
Education - primary			reference level		
Education - secondary	0.118***	0.118***	0.123***	0.118***	0.112***
Education - tertiary	0.466***	0.466***	0.475***	0.466***	0.459***
English - does not speak			reference level		
English - speaks not well	0.035***	.035***	0.038***	0.037***	0.011
English - speaks well	0.096***	0.095***	0.099***	0.100***	0.054***
English - speaks very well	0.286***	0.285***	0.286***	0.291***	0.228***
English - speaks only English	0.327***	0.326***	0.321***	0.332***	0.253***
Inv. Mill's Ratio	0.464***	0.464***	0.499***	0.463***	0.458***
Constant	-0.911***	-0.911***	-0.939***	-0.922***	-0.831***
Ethnic group dummies	no	no	no	no	no
Super-PUMA dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
No of observations	734,655	734,618	734,655	734,655	734,655
R^2	0.11	0.11	0.11	0.11	0.11

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

In Models 5-5c the overall effect of ethnic competition on business income is positive (lower value of HHI_{ethnic}) implies stronger competition). In general the length of stay is positively related to higher returns to self-employment. When interacted with the measure of ethnic competition, though, the duration appears to be negatively linked to profits. Among long-stayers HHI_{ethnic} has an increased negative relation to profits $(-0.009 + (-0.030) = -0.039$, compared to -0.009 among short-stayers). This leads us to hypothesize that ethnic enclaves may serve as business incubators. Members of dominant groups generally have higher profits. When we consider the way dominance interacts with ethnic competition, though, it turns out that being a member of a dominant ethnic group within an industry reduces the detrimental effect of ethnic competition on one's profits, but nevertheless the value of the estimator remains negative $(0.013 + (-0.016) = -0.003)$. For immigrants (non-citizens) ethnic competition decreases profits $(-0.050 + 0.057 = 0.007)$.

6.5 Summary and conclusions

The objective of this analysis was to test the relation between returns to business operation and co-ethnic complementarity and competition among Puerto Ricans in the US. Based on the findings of previous chapters of this thesis it was hypothesized that it is particularly the complementarity effects that provide an ultimately favorable environment for the development of ethnic and immigrant self-employment. Competition was expected to have adverse effects on the returns to immigrant self-employment.

Complementarity was operationalized as the relative number and wealth of Puerto Ricans living in an area. Competition was measured by means of the Hirschman-Herfindahl Index of market concentration calculated for a given industry and locality (general competition), and for Puerto Rican businesses operating in that industry and locality (ethnic competition). Mincerian wage equations were adapted to suit the research question and the Heckman Correction was employed to correct for selection bias.

The obtained results indicate that:

- overall market competition itself has no significant impact the profitability of Puerto Rican business;
- when ethnic competition is controlled for, though, we find that as the general level of competition in an industry increases (HHI falls), returns to business operation also increase, what implies that the general level of competition in an industry has a rather benign relation with Puerto Rican business profitability;
- nevertheless ethnic competition is negatively correlated with self-employment incomes of Puerto Ricans, as standard industrial organization theory would predict;
- the number of Puerto Ricans living in a given locality is generally also negatively associated with self-employment incomes of co-ethnics; this relation becomes insignificant, though, when ethnic competition is accounted for;
- the average wealth of the Puerto Rican community residing in a given area is unambiguously positively related to Puerto Rican business profitability;
- among standard covariates which supposedly affect one's productivity only age (non-linearly) and tertiary education (as compared to incomplete education) have been found to have a significant relation to business returns among the Puerto Ricans residing in the US;
- interestingly no selection bias has been found among the Puerto Ricans in terms their propensity to work rather than be (officially) inactive on the labor market;
- whether one is a Puerto Rican "immigrant" or "native" operating a business in the US is not significantly related to his/her business income.

From the first three results reported above we derive a conclusion that businesses run by Puerto Ricans are on one hand isolated (independent) from the industrial environment

in the local area (Models (1)-(3)), but once competition among Puerto Ricans themselves is accounted for (Model (4)), the local industrial relations have a beneficial effect on their profits - the more competition there is outside of the "enclave", the more money the enclave makes. What is more, competition among Puerto Ricans themselves follows a quite standard pattern in terms of its relation with business profitability. The fact that the greater the number of co-ethnics the lower the average business incomes (Model 2) could also be related to the ethnic competition. Yet another explanation can be found in relatively lower purchasing parity in large clusters of ethnic minorities. In order to control for this scenario we have have accounted for the relative wealth of co-ethnics (Model 3). This lead to concluding that indeed the richer the local Puerto Rican community, the higher the returns to business operation. This is an effect attributed to increased purchasing power parity, though it could indicate easier access to startup capital for larger (an possibly more profitable) businesses as well. We are unable to single out this feature, though.

From Model 4 we have concluded that when both competition and complementarity affects are accounted for, competition is the dominant force, having a twice as strong relation to business income. 'Monopolists', i.e. individuals who were the only Puerto Ricans running a business in a given industry and locality were partly responsible for this effect. After excluding observations for which the value of HHI or HHI_{PR} was 1 (Model 5) the estimator of general competition turned insignificant, but ethnic competition gained on value. The latter feature implies that the 'monopolists' had relatively low incomes, as they weakened the profit-concentration relationship. The reasons for such a situation may be sought after in the nature of those businesses which could have been a poor decision in the first place, what resulted in only one entrepreneur undertaking a given activity. Whether one is a Puerto Rican "immigrant" or was born in the US was not related to significantly different incomes.

The results obtained for Puerto Rican immigrants in the US are indicative of them functioning on integrated internal markets. The level of general competition is preferable in terms of higher returns to entrepreneurship. The level of intra-group competition had a, theoretically expectable, adverse relation to business profits. This led to thinking that these effects are in line with Puerto Ricans operating within ethnic enclaves. Moreover, rather saturated enclaves, at least market-wise. Ethnic enclaves, as described in detail in the introductory literature review to this thesis, are a form of protected markets. It has not been speculated what this isolation may indicate. The fact that increasing external (out-of-enclave) competition is beneficial to within-enclave business profits seems to be one of the mechanisms resulting from this isolation.

External validity for the Puerto Rican case study was provided through an identical analysis but run on eight aggregate ethnic groups living in the US. The results are indicative of them also being independent of the mainstream economy. Yet, they point to a quite different mechanism of how this isolation may affect returns to business activity.

Contrary to the result obtained for Puerto Ricans, the level of general competition has no significant effect on business profitability of the studied ethnic groups' businesses. When it comes to ethnic competition the result also differed from that obtained for Puerto Ricans. The analysis revealed that the higher the level of ethnic competition, the more favorable it is to ethnic entrepreneurs.

The insignificant relation of general competition to profits leads to concluding that ethnic groups (at an aggregate level, on average) also function in protected, or isolated markets, just as it was the case for Puerto Ricans. Given the fact that internal competition, rather than being detrimental, induces profits, we are also inclined to think that these ethnic markets are not saturated. This may be the second mechanism related to the isolable nature of ethnic markets.

Further inquiry into the mechanisms of how ethnic niches interact the wider economy, two of which have been recognized in this study, provide a fertile ground for future research.

Among the obvious data limitations related to aggregating the ethnic groups by approximately continental lines or using self-reported incomes as explanatory variables, one major assumption should be specifically stressed. In this study it has been assumed that people of a given ancestry actually operate "ethnic" businesses, i.e. that they capitalize their cultural capital in the course of their entrepreneurial activities. This, of course, is a very strong assumption. With the available data it is impossible, however, to determine the service or product profile of considered business owners. Nevertheless, it would be an equally strong assumption to neglect the common ethnic background of entrepreneurs as a form of monetizable cultural capital. This especially contradicts the assumptions underlying available theoretical concepts related to the economics of entrepreneurship, which heavily relay on the fact that we do not live in ethnicity-blind societies, but that rather the contrary is true.

Chapter 7

Summary and Conclusions

We live in a divided world. International borders separate the developed from the developing, the abundant from the scarce, the high earners from the low earners. Being on the disadvantageous side of the world's political, but also economic and social, frontiers may be limiting. The possibility to cross those frontiers, though, may be related to opportunities inexistent otherwise. This thesis explored whether immigrant self-employment can be a way of taking advantage of such opportunities.

The current state of the art in research on either migration, entrepreneurship or, specifically, immigrant entrepreneurship does not provide an answer to very basic questions concerning the choice of self-employment among international migrants. First, do self-employed migrants actually earn more than what they would have earned had they not migrated? Second, in light of the answer to the previous question, can self-employment be a way for migrants to deal with internal labor market imperfections at the destination country, like wage- and occupational-discrimination? Third, do increasing concentrations of immigrants hamper or help immigrant entrepreneurship?

To the best of the author's knowledge an answer to the first question has not been provided in existing literature. Thus we simply do not know whether, in general terms, immigrant self-employment should be treated as a story of economic success or a story of subsistence. Discrimination, in spite of having been examined among immigrants in general, has been rarely looked into as a feature affecting immigrant self-employment rates. Market competition and complementarity are studied in relation to the mainstream economy, but are under-researched notions, especially when considered jointly, in the context of ethnic entrepreneurship. Ethnic concentrations are known for boosting immigrant entrepreneurship by enabling employment of co-ethnic workers, increasing the pool of consumers, and providing access to startup capital. It has not been analyzed, though, to what extent one's co-ethnics could also be the source of inferior market competition.

Finding answers to the above questions is important for at least two reasons. On the academic level it significantly contributes to existing literature, filling in knowledge gaps which limit our understanding of the surrounding world. This is especially important in the age of dynamic inter-regional mobility, when majorities become minorities, when the crisis of multiculturalism is becoming a more and more recognizable issue, and when,

despite rising levels of human development, societies experience increasing inequalities in income, unsustainable patterns of consumption, and low social cohesion (United Nations Development Programme, 2013). Under the above circumstances finding answers to the questions posed by this thesis is also of practical importance. It provides insights into the nature of immigrant self-employment as labor market choice of individuals. Depending on whether immigrant entrepreneurship is rather that of "opportunity" or that of "necessity" makes it either an idea worth spreading or just a negative by-product of labor market imperfections existing across and within countries.

On the basis of existing evidence this thesis hypothesized that immigrant self-employment is a rather beneficial undertaking. The general hypothesis stated that

General hypothesis *Self-employment, on top of international labor migration, facilitates reducing the inefficiency of the international division of labor.*

Given the positive verification of the operational hypotheses formulated for the purpose of this thesis, the studies conducted within the scope this doctoral research affirm the general hypothesis stated above. From this perspective the main conclusions of this thesis are the following:

- by reallocating to markets where one's skills, abilities or knowledge are relatively scarce, or add to the diversity of supplied products or services, individuals may experience significant income gains. Though migration as such is a profitable strategy, by pursuing self-employment immigrants may obtain even higher returns. Notably, this result has been obtained in a context free from legal barriers to mobility;

- the profitability of immigrant self-employment may not only allow to overcome the labor market inefficiencies related to the international division of labor, but also to the internal market divide, especially affecting groups with limited capacity of action and intervention. In this sense it appears that, at least for some immigrant groups self-employment can actually be the first-best option in terms of returns to productivity. What is more, if small social distance between employees and employers is related to smaller discrimination, then the development of ethnic enterprises may not only be beneficial for the entrepreneurs themselves, but also for the co-ethnic workers they might potentially hire;

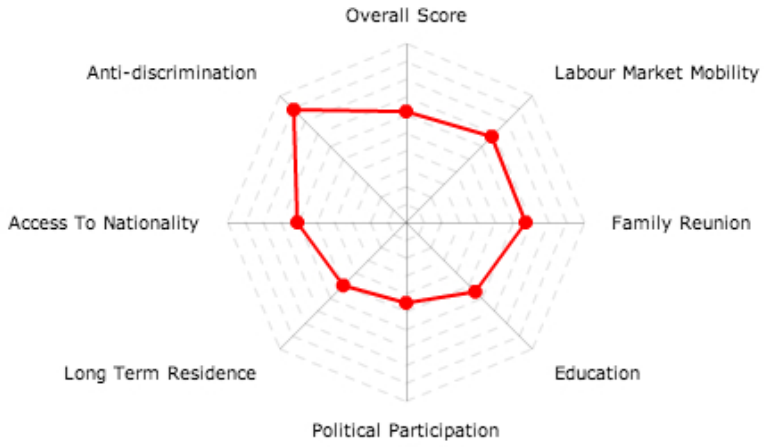
- the multilateral benefits of immigrant self-employment can be obtained by juts letting it happen. By clustering and taking advantage of their cultural and social capital migrants may develop internal markets which could serve as natural incubators for businesses. Extreme expansion of these markets need not to be a worry for native/majority firms, though, as it appears that ethnic economies are internally subject to general market rules determining the relation of competition to profits.

The above-mentioned conclusions can be put in the perspective of actual US immigration policies as evaluated by MIPEX, the Migrant Integration Policy Index. MIPEX

comprises 148 policy indicators on migrant integration. They are designed to benchmark current laws and policies against the highest standards formulated by European Union's or international organizations' directives, conventions, etc. MIPEX covers seven policy areas of integration: labor market mobility, family reunion, access to education, political participation, long-term residence, access to citizenship, and anti-discrimination laws and protections. "Anti-discrimination laws and protections", "labor market mobility", and "long-term residence" are especially important for labor market analyses.

As far as anti-discrimination policies are concerned, the US scored highest among 31 countries considered (89/100), Figure 7.1. Its efforts in terms of enabling immigrant labor market mobility (including recognition of qualifications) or long term residence, are much weaker, though (68/100 and 50/100 respectively). It is the latter two aspects, however, that in light of the conducted studies reveal greatest potential in terms of facilitating immigrant incorporation into the receiving markets. Needless to say at this point, this can be done very effectively by enabling immigrant self-employment.

Figure 7.1: Migrant Integration Policy Index scores for the US, by policy areas



Source: www.mipex.eu/usa.

Chapter 5 of this thesis concluded that immigrant labor market discrimination in the US exists and fares well. Across all considered groups of immigrants we found significant and sizable bias in terms of their pay. Of course, the applied measure could be imperfect, but the results are fairly consistent. Thus it seems that in practice combating discrimination works rather poorly.

One might think that investing in anti-discrimination measures and in their application can not be a bad idea, as without these measures the immigrants' situation could have been even worse. But since, under certain circumstances, discrimination induces immigrant self-employment then maybe it is worthwhile to consider creating these preferential circumstances rather than taking action against discrimination as such? A circumstance which emerged as critical for increasing the intensity of immigrant entrepreneurship under discrimination was the relative profitability of self-employment over wage-employment (as demonstrated in Chapters 4 and 5). Thus creating and nurturing environments which

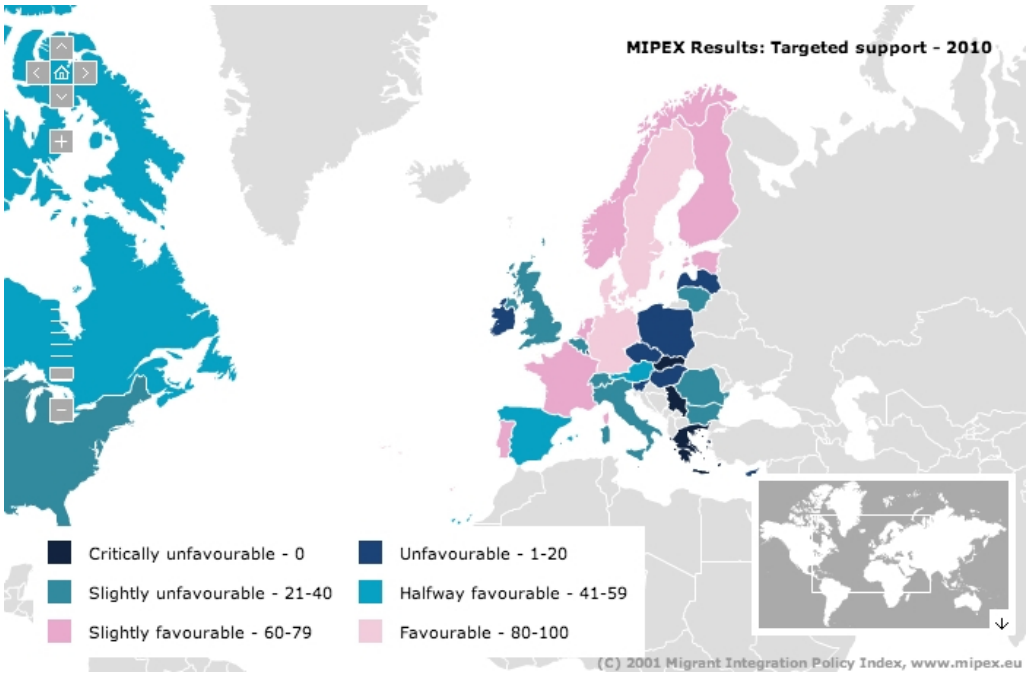
support business creation and development might be a very efficient way of enabling immigrant labor market integration.

When exploring the returns to immigrant entrepreneurship in Chapter 6 we found that among factors which are positively related to the profitability of immigrant businesses there is ethnic congestion (especially associated with long stay), tertiary education, and possibly also the level of English language command. These issues are addressed by two policy areas in which, according to MIPEX, the US performs relatively poorly: labor market mobility and long-term residence.

The "labor market mobility" policy area comprises four composite indicators: access, access to general support, targeted support, and workers' rights. The US scores at the top in terms of ensuring that resident immigrants and their families have similar chances in the labour market as native-born Americans. It is emphasized, though, that because some states and professional organizations are not co-operating to recognize foreign diplomas, the jobs immigrants find may be far below their qualifications. As troubling as the latter may seem, on one hand, Chapter 5 showed that even if immigrants' skills were rewarded like those of natives, they would still earn relatively less. On the other hand, Chapter 6 revealed that higher education is related to higher returns to self-employment, but recall that in case of the self-employed their formal education does not need formal recognition. Thus immigrant entrepreneurship, by employing immigrants' skills as they are rather than as how they are formally recognized, overcomes the imperfections of how foreign education is (not) recognized.

An aspect of "labor market mobility" policies which should be more troubling, though, is that the US score 25/100 on "targeted support", placing it at the lower half of the MIPEX ranking, Figure 7.2.

Figure 7.2: MIPEX "Targeted support" policy scores



Source: www.mipex.eu/usa.

Throughout this thesis it has been emphasized and proven that immigrants are a very specific type of labor force, one which is able to develop its fullest potential once it can preserve its uniqueness. This is most visible in the operation of ethnic economies where entrepreneurs can benefit from the presence of other ethnic entrepreneurs and of co-ethnic pools of consumers and workers, as demonstrated in Chapter 6. The same chapter also pointed to the fact that in some cases (visible in the extension but not in the Puerto Rican case), improved knowledge of English may also be very important for higher business returns. The one thing ethnic economies and space for language proficiency have in common is that they are both inherently related to immigrants, as opposed to natives. Thus integration strategies which support newcomers in establishing businesses by targeting their specific needs, such as language training could be most favorable. Nevertheless, even here, the policy does not have to be general, as for some groups the mere ability to become self-employed is as much as they need to succeed.

Last but not least, the conducted research pointed to duration of stay as a policy-dependent factor which is related to higher returns to business operation (Chapter 6). This connection may arise from the possibilities and perspectives for investment, as well as from the timely acquisition of local cultural and social capital at the destination. MIPEX gave the US one of the lowest scores (50/100) on the "long-term residence" measure due to the country's complex immigration laws, limited visa availability, high fees, and long processing queues, which make immigrants' socio-economic integration challenging. The possibly detrimental effects of migration under the circumstance of short business operation have also been signaled in Chapter 4. The same chapter also provided an insight as to the possibly high returns obtained under no legal barriers to entry. Further studies on this issue will have to be performed, though, in order to enable more conclusive statements.

Summarizing, it seems that as far as economic outcomes of immigrants are concerned self-employment and its facilitation at the destination countries may be a remedy not only for existing labor market imperfections, but also for the deficiencies of immigration policies. Rather than politically creating an ethnicity-blind market in which everyone is granted equal opportunities, it might more efficient to enable the market to recognize existing differences and allow people to take advantage of them.

The conclusions and policy recommendations presented above have been developed based on empirical tests of three operational hypotheses. Respective analyses have been organized into three subsequent chapters of the thesis, preceded by a review of existing theoretical frameworks useful for the analyses.

Chapter 4 was an analysis of the income gains to immigrant self-employment. There are two main contributions this chapter made to the existing state of the art. First, the conducted research took advantage of Propensity Score Matching (PSM) as a method for establishing a counterfactual for the group of self-employed migrants. This technique has been rarely used in studies on the economics of migration while it provides a neat

solution to the problem of immigrant selectivity present in all empirical analyses. On one hand, the conducted research thus widens the array of possible applications of PSM and, on the other hand, provides migration analysts with guidance as to the applicability of PSM in their studies. Second, the analysis found that immigrant self-employment may indeed be an income-maximizing strategy. The study reached this conclusion based on a number of comparisons concerning varying labor market states and both the home and host labor markets. The study assessed wage-employment and self-employment at home as well as wage-employment abroad as possible alternatives. Having established that immigrant self-employment is a profitable choice in comparison to the alternatives, the next chapter searched for the answer to the question of whether the incidence of immigrant self-employment is related to labor market discrimination at the destination. Thereby it aimed at determining if immigrant self-employment is an answer to labor market imperfections at the destination.

Chapter 5 involved a two-stage procedure. First, the Oaxaca-Blinder decomposition of wages was carried out for immigrants and natives. This allowed to determine the extent of labor market discrimination of immigrants across US states. Second, this measure was used in models estimating the probability of immigrant self-employment. Given the fact that this chapter did not aim at measuring discrimination *per se*, but rather discrimination as a factor considered in a further decision-making process, two moderating circumstances have been considered: labor market duality and relative deprivation. The former was incorporated into the analysis by performing comparisons of estimates obtained from models which included or excluded occupation as a determinant of wages. The latter was considered in that labor market discrimination was computed by comparing the earnings of immigrants to the US with the earnings of several possible reference groups of natives.

The notion of employing several comparison groups, rather than just one, was a pioneering approach in the study of immigrant wage discrimination. Usually only one comparison group is considered and no distinction is made between the social and economic significance of discriminatory practices. As this study showed, such a differentiation might yield quite new insights into the nature of discrimination. The primary finding and contribution to existing literature made by this chapter was that indeed the extent of discrimination differs depending on comparison group. Significant regional differences in the extent of labor market discrimination of immigrants have also been identified.

As far the relation of discrimination to the intensity of immigrant self-employment is concerned the obtained results proved that if self-employment is a more profitable alternative to wage-employment higher levels of wage-discrimination are related to greater concentrations of immigrant self-employment. This result was quite unambiguous in terms of which group of natives was considered as the reference. This outcome points to the rather gainful role of immigrant self-employment in reducing the adverse effects of labor market imperfections, be it across or within markets. At this stage it remained a question, though, what are the limits of these gains.

In search for an answer to the question of how limited are the benefits of immigrant self-employment, Chapter 6 looked into the notion of ethnic competition. The intuition behind this study was that since self-employment can be such a beneficial strategy it may generate a crowding-in effect which would eventually become detrimental to profits. In order to measure the extent of competition the Hirshman-Herfindahl Index of industrial concentration (HHI) has been computed for ethnic groups across local areas in the US. This, so called, ethnic HHI, together with a measure of the local concentrations of these ethnic communities (ratio of the number of co-ethnics to the number of non co-ethnics) was used in estimating the returns to ethnic entrepreneurship. In this manner the study dealt not only with the competition effects, but also with ethnic complementaries.

Though the notion that ethnic concentrations can be beneficial for business has been already recognized in the literature, ethnic competition, especially that related to the functioning of ethnic businesses has not been studied. On that account the research refined the understanding of how ethnic businesses operate what is a significant contribution to the existing state of the art. The results indicate that ethnic communities indeed operate within sheltered markets, which can play the role of business incubators. In this way immigrants can capitalize on their ethnicity.

All three aspects of self-employment studied in this thesis are very clearly attributable to being an immigrant. The specificity of the gain experienced due to crossing an international border is quite straightforward. The notion of discrimination arises from the fact of being a minority population - a feature inexistent in the immigrants' country of origin. Ethnic competition and complementarity, despite not being analytically limited to immigrants *per se*, are also identifiable only in a multicultural context. And despite the fact that not all ethnic businesses must be immigrant businesses, all immigrant businesses are to at least some extent ethnic in nature.

The geographic scope of empirical inquiries included in this thesis has been initially limited to Puerto Rican immigration to the US. This was a choice driven both methodologically and conceptually. The US is the largest immigration country in the world, with 42 million residents born abroad (United Nations Population Division, 2010). That is roughly 14% of the total population. Having even a small sample of such a population enables quantitative analyses of the very small and specific subpopulation of self-employed immigrants. On top of the above condition the data had to enable estimations of returns to entrepreneurial activity. The US is one of the few countries which collects and provides public access to such data.

The primary reason for choosing Puerto Rico as the source country for the analysis was that it provides data which is consistent with that collected in the US, i.e. it also includes a variable reporting one's returns to self-employment. Moreover, Puerto Ricans do not experience any formal labor market restrictions in the US. This latter fact was critical for being able to neglect the issue of administratively-driven selectivity of migration. The former issue, i.e. having data on earnings for self-employment for both the sending and receiving country was vital for the estimations of the gains to immigrant entrepreneurship.

Being aware of the interpretational limitations of studies carried out in one specific context each chapter also comprised an extension, the aim of which was to provide external validity for the results obtained for Puerto Rican immigrants. The extension to Chapter 4 estimated the returns to self-employment among inter-state migrants in the US. The extension to Chapter 5 dealt with labor market discrimination and the intensity of self-employment among eight aggregate ethnic groups in the US, delimited approximately by continental boundaries. The same ethnic groups were also considered for the purpose of providing external validity for the study in Chapter 6.

The empirical studies conducted within the framework of this thesis are summarized in Table 7.1.

Table 7.1: Summary of conducted empirical research

	Chapter 4: Gains to immigrant self-employment	Chapter 5: Discrimination	Chapter 6: Competition and complementarity
theoretical foundation	neoclassical migration theory; New Economics of Labor Migration	discrimination/disadvantage theory; employer, consumer and credit market discrimination	ethnic enclave hypothesis, theory of industrial organization
identified research gap	no study on the gains to both migration and self-employment considered jointly; no study concerning opportunity costs related to various forms of labor market activity; only one application of the proposed method	lack of research considering various reference groups in the analysis of wage discrimination; few studies exploiting discrimination as a decision-making factor in the pursuit of self-employment among immigrants; no study analyzing both wage and occupational discrimination in both hourly and yearly terms	lack of research on the effects of business competition among co-ethnic entrepreneurs; no studies no the joint effect of ethnic competition and complementarity on local markets
research question	do self-employed immigrants earn more than what they would if they had not migrated?	how is labor market discrimination related to the incidence of immigrant self-employment?	how do ethnic competition and ethnic complementarity relate to returns to ethnic entrepreneurship?
research hypothesis	Income from self-employment abroad is greater than expected income earned at home	Labor market discrimination is related to increased intensity of immigrant self-employment	Ethnic competition is related to lower, while ethnic complementarity is related to higher returns to local entrepreneurial activity
methods/measures	Propensity Score Matching	Oaxaca-Blinder wage decomposition, probit model, Propensity Score Matching	Ordinary Least Squares, Hirschman-Herfindahl Index
external validity	US inter-state migration	all aggregate US immigrant ethnic groups	all aggregate US ethnic groups

	Chapter 4: Gains to immigrant self-employment	Chapter 5: Discrimination	Chapter 6: Competition and complementarity
contribution	finding that immigrant self-employment may be a more profitable choice than that of employment at home or even wage-employment abroad; Propensity Score Matching is an applicable method to the study of income gains from migration	discovering that the extent of discrimination varies across local markets and comparison groups; if self-employment is a relatively more profitable form of labor market activity than wage-employment, then greater wage discrimination will be related to increased intensity of immigrant self-employment	proving that migrants operate in sheltered markets which are subject to internal competition, but which operate independently of external competition
further inquiry	considering the possible self-deprivation of immigrants abroad, i.e. decreased standards of living; accounting for imperfect transferability of skills; conducting the study in other geographic contexts; determining what features of local markets at the destination are responsible for the extent of the gains	analysis of discrimination in majority-minority contexts; studies on social distance and relevant comparisons as labor market decision-making factors; research on internal mobility related to avoiding wage discrimination as an alternative to self-employment	analysis of ethnic competition in majority-minority contexts; considering the possibility of ethnic market saturation and its relation to ethnic business profitability

The results reported above are valid only under the assumptions underlaying each of the conducted studies. First, they have been developed on the basis of analyses of declarative data including one's earnings, labor market status or ethnic origin. Nevertheless, it has been assumed to be the best available approximation of the actual state of things. A more reliable and at least equally rich data source has not been identified. Second, the considered research methods are applicable only under specific assumptions. It has been reasoned that these assumptions have been met in each case and that the interpretation of the results is valid as long as they are fulfilled. Third, among the interpretational limitations it should also be stressed that the research considered a very specific context of immigration to the US. The external validity of the conducted studies suggested that certain aspects of immigrant self-employment are rather general (discrimination, ethnic market isolation), while others are highly context specific (income gains, relation of discrimination to self-employment intensity). Any generalizations of the findings should thus account for the recognized similarities and differences.

In light of these limitations there is a great potential for future research, though. As already indicated in Table 7.1, the possible areas of further inquiry comprise, but are not limited to:

- researching how self-deprivation and self-exploitation may increase the gains to migration and to immigrant self-employment and whether these are viable strategies e.g. in terms of immigrants health in the long run;

- accounting for the imperfect transferability of skills between the countries of origin and destination and whether self-employment, contrary to wage-employment, can reduce the adverse effects of these imperfections;
- exploring how wage discrimination and ethnic competition and complementarity are related to labor market outcomes of immigrants in majority-minority areas; and how such a context affects the natives - their labor market choices and returns to economics activity;
- studying on a more in-depth level how immigrants construct their reference groups and how their interrelations affect immigrants' choices;
- mapping international migrants' internal migration at the destination in search for the optimal location for their businesses;
- analyzing ethnic market saturation and its effect on the profitability and organization of businesses within ethnic enclaves.

It seems that, despite the forces of globalization and market organization there is and, at least for some time, still will be "structural viability" of small-businesses in general (Kloosterman and Rath, 2001) and, thus, also potential for immigrant ownership of these small businesses. Demand exerted by consumers employed in large-scale enterprises for group-specific, nearly tailor-made products and services in which it is hard to achieve economies of scale (child-care, house-cleaning etc.) is precisely where immigrant entrepreneurs often find their niche. As migration is a process which can not be stopped, but at most managed, this thesis provided guidance as to whether and how migrants can benefit for self-employment as a labor market strategy, leaving many more fascinating avenues open to further academic inquiry.

List of Figures

List of Tables

Bibliography

Amin, A. (1999, 06). An Institutionalist Perspective on Regional Economic Development. *International Journal of Urban and Regional Research 23*(2), 365–378.

Anacka, M. and M. Okólski (2010). *A Continent Moving West? EU Enlargement and Labour Migration from Central and Eastern Europe*, Chapter Direct demographic consequences of post-accession migration for Poland, pp. 141–163. Amsterdam University Press.

Anderson, L. (2011). Occupational choice and returns to self-employment among immigrants. *International Journal of Manpower 32*(8), 900–922.

Arrow, K. J. (1998). What Has Economics to Say About Racial Discrimination? *The Journal of Economic Perspectives 12*(2), 91–100.

Åstebro, T. and J. Chen (2012). The Returns to Entrepreneurship: Measurements, Methods, Old and New Results. http://ssrn.com/abstract=2125887.

Audretsch, D., M. Carree, and A. Thurik (2001, August). Does Entrepreneurship reduce Unemployment? Tinbergen Institute Discussion Papers 01-074/3, Tinbergen Institute.

Auster, E. and H. Aldrich (1984). Small business vulnerability, ethnic enclaves and ethnic entreprise. In R. Ward and R. Jenkins (Eds.), *Ethnic communities in business. Strategies for economic survival*, pp. 39–54. Cambridge University Press.

Bain, J. (1956). *Barriers to New Competition*. Harvard University Press.

Baptista, R. and R. Thurik (2004, August). The Relationship between Entrepreneurship and Unemployment: is Portugal an Outlier? Papers on Entrepreneurship, Growth and Public Policy 2004-40, Max Planck Institute of Economics, Entrepreneurship, Growth and Public Policy Group.

Beaujot, R., P. S. Maxim, and J. Z. Zhao (1994). Self-employment Among Immigrants: A Test of the Blocked Mobility Hypothesis. *Canadian Studies in Population 21*(2), 81–96.

Becker, G. S. (1971). *The Economics of Discrimination*. Economic Research Studies. University of Chicago Press.

Becker, S. O. and A. Ichino (2002). Estimation of average treatment effects based on propensity scores. *The Stata Journal 2*(4), 358–377.

Berglann, H., E. Moen, K. Røed, and J. Skogstrøm (2011). Entrepreneurship: Origins and returns. *Labour Economics 18*(2), 180–193.

Bhagwati, N. and P. Srinivasan (1998). *Lectures on International Trade*. The MIT Press.

Blanchflower, D. G. and A. J. Oswald (1998, January). What Makes an Entrepreneur? *Journal of Labor Economics 16*(1), 26–60.

Blanchflower, D. G. and A. J. Oswald (2007, November). What Makes a Young Entrepreneur. IZA Discussion Papers 3139, Institute for the Study of Labor (IZA).

Bodvarsson, O. B. and H. van der Berg (2009). *The Economics of Immigration. Theory and Policy*. Springer-Verlag.

Bonacich, E. (1973). A Theory of Middleman Minorities. *American Sociological Review 38*, 583–594.

Borjas, G. (1987, September). Self-Selection and the Earnings of Immigrants. *The American Economic Review 77*(4), 531–553.

Borjas, G. (1988). Self-Selection and the Earnings of Immigrants. Working Paper 2248, National Bureau of Economic Research.

Borjas, G. (1990). *Friends or Strangers, the Impact of Immigrants on the U.S. Economy*. Basic Books.

Borjas, G. (2000). *Issues in the Economics of Migration*. University of Chicago Press.

Borjas, G. J. (1986). The Self-Employment Experience of Immigrants. *The Journal of Human Resources 21*(4), 485–506.

Borjas, G. J. and S. G. Bronars (1989, September). Consumer Discrimination and Self-Employment. NBER Working Papers 2627, National Bureau of Economic Research, Inc.

Brock, W. A. and D. S. Evans (1986). *The Economics of Small Businesses: Their Role and Regulation in the US Economy*. New York: Holmes&Meier.

Caliendo, M. and S. Kopeinig (2005). Some Practical Guidance for the Implementation of Propensity Score Matching. IZA Discussion Papers 1588, Institute for the Study of Labor (IZA).

Caliendo, M. and A. S. Kritikos (2008). Is Entrepreneurial Success Predictable? An Ex-Ante Analysis of the Character-Based Approach. *Kyklos 61*, 189 – 214.

Caliendo, M. and A. S. Kritikos (2012). Searching for the Entrepreneurial Personality: New Evidence and Avenues for Further Research. *Journal of Economic Psychology 33*(2), 319–324.

Cameron, A. and P. Trivedi (2009). *Microeconometrics using stata*, Volume 5. Stata Press College Station, TX.

Cantillon, R. (1755). Essai sur la nature du commerce en général. Traduit de l'anglois.

Carlton, D. W. and J. M. Perloff (2005). *Modern Industrial Organization* (Fourth ed.). Pearson Addison Wesley.

Castles, S. and M. J. Miller (2009). *The Age of Migration*. The Guilford Press.

Central Intelligence Agency (2012). *The World Factbook*. Central Intelligence Agency.

Clark, K. and S. Drinkwater (1998, August). Ethnicity and Self-Employment in Britain. *Oxford Bulletin of Economics and Statistics 60*(3), 383–407.

Clark, K. and S. Drinkwater (2000). Pushed uot or pulled in? Self-employment among thnic minoroties in England and Wales. *Labour Economics* (7), 603–628.

Clemens, M. (2010, June). The Roots of Global Wage Gaps: Evidence from Randomized Processing of U.S. Visas. Working Paper 212, Center for Global Development.

Clemens, M., C. Montenegro, and L. Pritchett (2009). The place premium: wage differences for identical workers across the US border. *Harvard Kennedy School Faculty Research Working Paper Series RWP09-004*.

Coate, S. and S. Tennyson (1992, April). Market Discrimination, Imperfect Information and Self-employment. *Oxford Economic Papers 44*(2), 272–288.

Cowling, K. and M. Waterson (1976, August). Price-Cost Margins and Market Structure. *Economica 43*(171), 267–274.

Cowling, M. and P. Mitchell (1997, September). The Evolution of U.K. Self-Employment: A Study of Government Policy and the Role of the Macroeconomy. *The Manchester School of Economic & Social Studies 65*(4), 427–42.

Curry, B. and K. D. George (1983, March). Industrial Concentration: A survey. *The Journal of Industrial Economics 31*(3), 203–255.

DeJong, G. F. (2000). Expectations, gender, and norms in migration decision-making. *Population studies 54*(3), 307–319.

Demsetz, H. (1973, April). Industry Structure, Market Rivalry and Public Policy. *Journal of Law and Economics 16*(1), 1–9.

Dickson, P. H., G. T. Solomon, and K. M. Weaver (2008). Entrepreneurial Selectiona And Success: Does Education Matter? *Journal of Small Business and Enterprise Development 15*(2), 239–258.

Dixit, A. and R. Rob (1994, February). Switching Costs and Sectoral Adjustments in General Equilibrium with Uninsured Risk. *Journal of Economic Theory 62*(1), 48–69.

Doeringer, P. B. and M. J. Piore (1971). *Internal labor markets and manpower analysis*. Heath Lexington Books.

Dolton, P. J. and G. H. Makepeace (1987). Sample Selection and the Male-Female Earnings Differential in the Graduate Labour Market. *Oxford Economic Papers 38*, 317–341.

Dumon, W. A. (1989, June). Family and Migration. *International Migration 27*(2), 251–270.

Earle, J. S. and Z. Sakova (2000, September). Business start-ups or disguised unemployment? Evidence on the character of self-employment from transition economies. *Labour Economics 7*(5), 575–601.

Eisenstadt, S. N. (1953). Analysis of Paterns of Immigration and Absorption of Immigrants. *Population studies 7*(2), 167–180.

Encyclopedia Britannica (2012). *Britannica - The Online Encyclopedia*. Encyclopedia Britannica Inc.

Evans, M. D. R. and J. Kelley (1991, November). Prejudice, Discrimination, and the Labor Market: Attainments of Immigrants in Australia. *American Journal of Sociology 97*(3), 721–759.

Fairlie, R. (2005, October). Entrepreneurship and Earnings among Young Adults from Disadvantaged Families. *Small Business Economics 25*(3), 223–236.

Fairlie, R. W. and B. D. Meyer (1996, Autumn). Ethnic and Racial Self-Employment Differences and Possible Explanations. *The Journal of Human Resources 31*(4), 757–793.

Ferrante, F. and F. Sabatini (2007, January). Education, social capital and entrepreneurial selection in Italy. MPRA Paper 2451, University Library of Munich, Germany.

Ginevicius, R. and S. Cirba (2007). Determining market concentration. *Journal of Business Economics and Management 8*(1), 3–10.

Ginevicius, R. and S. Cirba (2009). Additive Measurement of Market Concentration. *Journal of Business Economics and Management* (3), 191–198.

Glaeser, E. L. and D. C. Mare (2001, April). Cities and Skills. *Journal of Labor Economics 19*(2), 316–342.

Global Entrepreneurship Monitor (2013). Key Indicators and Definitions.

Hamilton, B. H. (2000). Does Entrepreneurship Pay? An Empirical Analysis of the Returns to SelfEmployment. *Journal of Political Economy 108*(3), 604–631.

Hammarstedt, M. (2006). The Predicted Earnings Differential and Immigrant Self-employment in Sweden. *Applied Economics 38*(6), 619–630.

Harris, J. R. and M. P. Todaro (1970). Migration, unemployment and development: A two-sector analysis. *The American Economic Review 60*(1), 126–142.

Hartog, J., M. van Praag, and J. van der Sluis (2008, August). If You Are So Smart, Why Aren't You an Entrepreneur? Returns to Cognitive and Social Ability: Entrepreneurs versus Employees. IZA Discussion Papers 3648, Institute for the Study of Labor (IZA).

Hatton, T. J. and J. G. Williamson (2004, May). International Migration in the Long-Run: Positive Selection, Negative Selection and Policy. Working Paper 10529, National Bureau of Economic Research.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica 47*, 475–492.

Hicks, J. (1932). *The Theory of Wages*. Macmillan.

Huber, M., M. Lechner, and C. Wunch (2010, October). How to Control for Many Covariates? Reliable Estimators Based on the Propensity Score. Discussion Paper 5268, Institute for the Study of Labor (IZA).

Hughes, K. D. (2003). Pushed or Pulled? Womenís Entry into Self-Employment and Small Business Ownership. *Gender, Work and Organization 10*(4).

Hyytinen, A. and P. Ilmakunnas (2007, October). What distinguishes a serial entrepreneur? *Industrial and Corporate Change 16*(5), 793–821.

Jann, B. (2008). The blinder-Oaxaca decomposition for linear regression models. *The Stata Journal 8*(4), 453–479.

Jiobu, R. M. (1988, June). Ethnic Hegemony and the Japanese of California. *American Sociological Review 53*(3), 353–367.

Kennan, J. and J. R. Walker (2011, January). The Effect of Expected Income on Individual Migration Decisions. *Econometrica 79*(1), 211–251.

Kloosterman, R. (2004). Mixed Embeddedness as a Conceptual Framework for Exploring Immigrant Entrepreneurship. EUREX Lecture Nr. 8.

Kloosterman, R. and J. Rath (2001, April). Immigrant entrepreneurs in advanced economies: mixed embeddedness further explored. *Journal of Ethnic and Migration Studies 27*(2), 189–202.

Kofman, E. (2004). Familyrelated migration: a critial review of European Studies. *Journal of Ethnic and Migration Studies 30*(2), 243–262.

Le, A. T. (1999, September). Empirical Studies of Self-Employment. *Journal of Economic Surveys 13*(4), 381–416.

Lee, E. (1966). A theory of migration. *Demography 3*, 47–57.

Lee, S. Y., R. Florida, and Z. Acs (2004, January). Creativity and Entrepreneurship: A Regional Analysis of New Firm Formation. *Regional Studies 38*(8), 879–891.

Lehmer, F. and J. Ludsteck (2011, Winter). The Immigrant Wage Gap in Germany: Are East Europeans Worse Off? *International Migration Review 45*(4), 872–906.

Leibenstein, H. (1968). Entrepreneurship and Development. *American Economic Review 58*(2), 72–83.

Levine, R. and Y. Rubinstein. Does Entrepreneurship Pay? The Michael Bloombergs, the Hot Dog Vendors, and the Returns to Self-Employment.

Li, P. S. (2000, Winter). Economic Returns of Immigrants' Self-employment. *The Canadian Journal of Sociology 25*(1), 1–34.

Light, I. (1972). *Ethnic enterprise in America: business and welfare among Chinese, Japanese, and Blacks*. University of California Press.

Light, I. (2004). The Ethnic Economy. In N. Smelser and R. Swedberg (Eds.), *Handbook of Economic Sociology* (2nd ed.)., Chapter 28. Russell Sage Foundation.

Light, I., G. Sabagh, M. Bozorgmehr, and C. Der-Martirosian (1994, February). Beyond the Ethnic Enclave Economy. *Social Problems 41*(1), 65–80.

Llisterri, J. J., H. Mantis, P. Angelelli, and L. Tejerina (2006). Is youth entrepreneurship a necessity or an opportunity? : a first exploration of households and new enterprises surveys in latin america. Sustainable Development Department Technical paper series MSM-131.

Lofstrom, M. (2002). Labor market assimilation and the self-employmnt decision of immigrant entrepreneurs. *Journal of Population Economics 15*, 83–114.

Lucas, R. E. (1978). On the size distribution of business firms. *The Bell Journal of Economics 9*(2), 508–523.

Malecki, E. J. (1993). Entrepreneurship in Regional and Local Development. *International Regional Science Review 16*(1-2), 119–153.

Massey, D. S. (1999). Why does immigration occur? a theoretical synthesis. In C. Hirschman, P. Kasinitz, and J. DeWind (Eds.), *Handbook of International Migration: The American Experience*. The Russell Sage Foundation.

Massey, D. S., M. J. Fischer, and C. Capoferro (2006, December). International Migration and Gender in Latin America: A Comparative Analysis. *International Migration 44*(5), 63–91.

Mata, F. and R. Pendakur (1999). Immigration, Labor Force Integration and the Pursuit of Self-Employment. *International Migration Review 33*(9), 378–402.

Maxim, P. S. (1992, May). Immigrants, Visible Minorities, and Self-Employment. *Demography 29*(2), 181–198.

McKenzie, D., J. Gibson, and S. Stillman (2010). How Important is Selection? Experimental vs. Non-Experimental Measures of the Income Gains from Migration. *Journal of the European Economic Association 8*(4), 913–945.

McQuaid, R. W. (2002, August). Entrepreneurship and regional development policies. ERSA conference papers ersa02p187, European Regional Science Association.

Mincer, J. (1977). Family Migration Decisions. NBER Working Papers 199, National Bureau of Economic Research.

Model, S. (1992, Spring). The Ethnic Economy: Cubans and Chinese Reconsidered. *The Sociological Quarterly 33*(1), 63–82.

Moore, R. L. (1983, August). Employer Discrimination: Evidence From Self-Employed Workers. *The Review of Economics and Statistics 65*(3), 496–501.

Naude, W. (2008). Entrepreneurship in Economic Development. Working Papers RP2008/20, World Institute for Development Economic Research (UNU-WIDER).

Nestorowicz, J. (2008). Do Networks do the Works? Indian Entrepreneurs in Poland and Belgium. Working Papers 38/(96), Centre of Migration Research.

Nielsen, H. S., M. Rosholm, N. Smith, and L. Husted (2001, September). Qualifications, Discrimination, or Assimilation? An Extended Framework for Analysing Immigrant Wage Gaps. Disscussion Paper Series 365, IZA.

Oyelere, R. U. and W. Belton (2012). Comming to America: Does havinga developed home country matter for self-employment in the U.S.? Technical report, Georgia Institute of Technology.

Parker, S. C. (2006). *The Economics of Self-Employment and Entrepreneurship*. Cambridge.

Peltzman, S. (1977, October). The Gains and Losses from Industrial Concentration. *Journal of Law and Economics 20*(2), 229–263.

Pessar, P. R. and S. J. Mahler (2003, September). Transnational Migration: Bringing Gender In. *International Migration Review 37*(3), 812–846.

Piore, M. J. (1979). *Birds of Passage. Migrant Labor in Industrial Societies*. Cambridge University Press.

Portes, A. and L. Jensen (1989, December). The Enclave and the Entrants: Patterns of Ethnic Enterprise in Miami before and after Mariel. *American Sociological Review 54*(6), 929–949.

Portes, A. and L. Jensen (1992, June). Disproving the Enclave Hypothesis: Reply. *American Sociological Review 57*(3), 418–420.

Portes, A. and R. G. Rumbaut (1996). *Immigrant America: a portrait* (2nd ed.). University of California Press.

Portes, A. and S. Shafer (2006). Revisitng the Enclave Hypothesis: Miami Twenty-Five Yeras Later. Technical report, Princeton University.

Portes, A. and M. Zhou (1996, April). Self-employment and the Earnings of Immigrants. *American Sociological Review 21*(2), 219–230.

Raijman, R. and M. Tienda (2000, Autumn). Immigrants' Pathways to Business Ownership: A Comparative Ethnic Perspective. *International Migration Review 34*(3), 682–706.

Rath, J. (2000). A game of ethnic musical Chairs? Immigrant businesses and the formation and succession of niches in the Amsterdam economy. In S. Body-Gendrot and M. Martinello (Eds.), *Minorities in European Cities. The Dynamics of Social Integration and Social Exclusionat the Neighborhood Level*. Macmillan Press.

Rath, J. (2001). Do Immigrant entrepreneurs Play the Game of Ethnic Musical Chairs? A Critique of Waldinger's Model of Immigrant Incporporation. In A. Messina (Ed.), *A Continuing Quandry for States and Societies: West European Imigration and Immigrant Policy in the New Century*. Greenwood Press.

Ravenstein, E. G. (1885). The laws of migration. *Journal of the Statistical Society of London 48*(2), 167–235.

Rees, H. and A. Shah (1986, January). An empirical analysis of self-employment in the u.k. *Journal of Applied Econometrics 1*(1), 95–108.

Rosenbaum, P. and D. B. Rubin (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika 70*(1), 41–55.

Roy, A. (1951, June). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers 3*(2), 135–146.

Ruggles, S., J. T. Alexander, K. Genadek, R. Goeken, M. B. Schroeder, and M. Sobek (2010). Integrated Public Use Microdata Series: Version 5.0. Machine-readable database, Minneapolis: University of Minnesota.

Sanders, J. M. and V. Nee (1987, December). Limits of Ethnic Solidarity in the Enclave Economy. *American Sociological Review 52*(6), 745–773.

Sanders, J. M. and V. Nee (1992, June). Problems in Respolving The Enclave Economy Debate. *American Sociological Review 57*(3), 415–418.

Say, J.-B. (1803). *A Treatise on Political Economy*. Philadelphia: Lippincott, Grambo Co.,.

Schumpeter, J. A. (1947, November). The Creative Response in Economic History. *The Journal of Economic History 7*(2), 149–159.

Silva, O. (2007, November). The Jack-of-All-Trades entrepreneur: Innate talent or acquired skill? *Economics Letters 97*(2), 118–123.

Simpson, W. and R. Sproule (1998). Econometric Analysis of Canadian Self-employment Using SLID. *Income and Labour Dynamics Working Paper Series 98-16*.

Sjaastad, L. A. (1962). The costs and returns of human migration. *The Journal of Political Economy 70*(5), 80–93.

Smith, J. A. and P. E. Todd (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics 125*, 305–353.

Stark, O. and D. E. Bloom (1985). The new economics of labor migration. *American Economic Review 75*(2), 173–178.

Stark, O. and R. E. Lucas (1985, October). Motivations to Remit: Evidence from Botswana. *Journal of Political Economy 93*(5), 901–918.

Storper, M. (1995). The Resurgence of Regional Economies, Ten Years Later: The Region as a Nexus of Untraded Interdependencies. *European Urban and Regional Studies 2*(3), 191–221.

Tamvada, J. P. (2007). Essays on Entrepreneurship and Economic Development. Dissertation, Georg-August-Universitat Gottingen.

Taylor, M. P. (1996, May). Earnings, independence or unemployment: Why become self-employed? *Oxford Bulletin of Economics and Statistics 58*(2), 253–66.

Tirole, J. (1988). *The Theory of Industrial Organization*. Massachusetts Institute of Technology.

Tyrowicz, J. and J. Nestorowicz (2011). Age and Entrepreneurship During Transition. *Ekonomista* (3), 371–395.

United Nations (2011). The Age and Sex of Migrants.

United Nations Development Programme (2013). *Human Development Report 2013*. UNDP.

United Nations Population Division (2010). Trends in International Migrant Stock.

van der Sluis, J., M. van Praag, and A. van Witteloostuijn (2004, September). Comparing the Returns to Education for Entrepreneurs and Employees. Tinbergen Institute Discussion Papers 04-104/3, Tinbergen Institute.

Van Praag, C. M. and H. Van Ophem (1995). Determinants of Willingness and Opportunity to Start as an Entrepreneur. *Kyklos 48*(4), 513–40.

van Tubergen, F. (2005). Self-employment of immigrants: A cross-national study of 17 western societies. *Social Forces 84*, 709–732.

Waldinger, R. (1987). Post-Industrial New York Changing Ladders and Musical Chairs: Ethnicity and Opportunity in Post-Industrial New York. *Politics Society 15*, 369–401.

Waldinger, R. (1996). *Still the promised city?: African-Americans and new immigrants in postindustrial New York*. Harvard University Press.

Waldinger, R., H. Aldrich, and R. Ward (Eds.) (1990). *Ethnic Entrepreneurs. Immigrant business in Industrial Societies*, Volume 1 of *Sage Series on Race and Ethnic Relations*. Sage Publications.

Waldinger, R., R. Ward, and H. Aldrich (1985, November). Ethnic Businesses and Occupational Mobility in Advanced Societies. *Sociology 19*(4), 586–597.

Werbner, P. (2001). Metaphores of Spatiality and Networks in the Plural City: A Critique of the Ethnic Enclave Economy Debate. *Sociology 35*(3), 671–693.

Wilson, K. L. and W. A. Martin (1982, July). A Comparison of the Cuban and Balck Economies in Miami. *The American Journal of Sociology 88*(1), 135–160.

Wilson, K. L. and A. Portes (1980, September). An analysis of the Labor Market Experiences of Cubans in Miami. *The American Journal of Sociology 86*(22), 295–319.

Zhou, M. and J. R. Logan (1989, October). Returns on Human Capital in Ethic Enclaves: New York City's Chinatown. *American Sociological Review 54*(5), 809–820.

Appendices

Appendix to Chapter 4.

Independent variables used for propensity score estimation

Variable	SE_{PR}	SE_{US}	WE_{PR}	WE_{US}
age <20	0.13	0.24	0.15	0.09
	(0.33)	(0.43)	(0.36)	(0.29)
age 20-30	0.20	0.24	0.22	0.22
	(0.40)	(0.43)	(0.42)	(0.41)
age 30-40	0.24	0.23	0.25	0.27
	(0.43)	(0.42)	(0.43)	(0.44)
age 40-50	0.25	0.20	0.24	0.26
	(0.44)	(0.40)	(0.43)	(0.44)
age 50-60	0.17	0.09	0.13	0.16
	(0.38)	(0.28)	(0.34)	(0.37)
gender-male	0.75	0.51	0.48	0.68
	(0.43)	(0.50)	(0.50)	(0.47)
gender-female	0.25	0.49	0.52	0.32
	(0.43)	(0.50)	(0.50)	(0.47)
education-incomplete	0.01	0.00	0.01	0.01
	(0.10)	(0.07)	(0.09)	(0.12)
education-primary	0.18	0.07	0.06	0.10
	(0.38)	(0.25)	(0.24)	(0.29)
education-secondary	0.43	0.36	0.41	0.39
	(0.50)	(0.48)	(0.49)	(0.49)
education-tertiary	0.38	0.56	0.52	0.50
	(0.49)	(0.50)	(0.50)	(0.50)
married-no	0.45	0.53	0.47	0.37
	(0.50)	(0.50)	(0.50)	(0.48)
married-yes	0.55	0.47	0.53	0.63
	(0.50)	(0.50)	(0.50)	(0.48)
family size 1	0.21	0.16	0.21	0.20
	(0.41)	(0.37)	(0.41)	(0.40)
family size 2-3	0.46	0.45	0.45	0.45
	(0.50)	(0.50)	(0.50)	(0.50)
family size 4-5	0.28	0.33	0.29	0.29
	(0.45)	(0.47)	(0.45)	(0.46)
family size 5<	0.05	0.06	0.05	0.05
	(0.22)	(0.23)	(0.22)	(0.22)
	industries			
agriculture, forestry, fishing, hunting	0.10	0.02	0.01	0.05
	(0.30)	(0.13)	(0.10)	(0.22)
mining	0.00	0.00	0.00	0.00
	0.00	(0.03)	(0.02)	(0.02)
construction	0.21	0.07	0.04	0.16
	(0.41)	(0.26)	(0.19)	(0.37)
manufacturing	0.03	0.11	0.08	0.03
	(0.16)	(0.31)	(0.27)	(0.18)
transport/communications	0.05	0.05	0.07	0.06
	(0.21)	(0.23)	(0.26)	(0.23)
wholesale	0.02	0.03	0.02	0.03
	(0.12)	(0.17)	(0.13)	(0.17)
retail	0.18	0.18	0.09	0.16
	(0.38)	(0.39)	(0.29)	(0.37)
finance/insurance	0.03	0.05	0.04	0.05
	(0.16)	(0.22)	(0.20)	(0.23)
business/repair services	0.13	0.06	0.04	0.12

Independent variables used for propensity score estimation

Variable	SE_{PR}	SE_{US}	WE_{PR}	WE_{US}
	(0.34)	(0.23)	(0.20)	(0.33)
personal services	0.10	0.03	0.03	0.08
	(0.30)	(0.17)	(0.16)	(0.26)
entertainment	0.01	0.01	0.01	0.01
	(0.10)	(0.08)	(0.10)	(0.11)
professional services	0.15	0.28	0.39	0.25
	(0.36)	(0.45)	(0.49)	(0.43)
public administration	0.00	0.11	0.17	0.00
	0.00	(0.31)	(0.38)	0.00
military	0.00	0.00	0.01	0.00
	0.00	(0.04)	(0.12)	0.00
	occupations			
manager, professional	0.20	0.24	0.30	0.32
	(0.40)	(0.43)	(0.46)	(0.47)
technician, salesman, administrative worker	0.19	0.31	0.26	0.21
	(0.39)	(0.46)	(0.44)	(0.41)
service worker	0.16	0.19	0.22	0.16
	(0.36)	(0.39)	(0.42)	(0.37)
farmer etc.	0.10	0.02	0.02	0.05
	(0.30)	(0.15)	(0.13)	(0.21)
precision or repair worker, craftsman	0.20	0.09	0.07	0.14
	(0.40)	(0.28)	(0.25)	(0.35)
operator, fabricator, laborer	0.15	0.15	0.13	0.11
	(0.36)	(0.35)	(0.33)	(0.32)
military	0.00	0.00	0.01	0.00
	0.00	(0.04)	(0.11)	0.00
Observations	4,606	46,614	45,075	3,466

Standard deviations in parentheses.

Income gains to migration: Average Treatment Effects on the Treated (ATT)

Variable	Treated	Controls	Difference	PSM S.E.	T-stat
Self-employed in Puerto Rico vs Self-employed in U.S.					
Caliper Matching (1 s.d.)					
hourly income	18.82	11.77	7.06	0.63	11.18
yearly income	38,939	19,584	19,355	1,206	16.05
Caliper Matching (2 s.d.)					
hourly income	18.82	13.23	5.89	0.65	8.67
yearly income	38,939	22,227	16,712	1,219	13.71
Caliper Matching (0.5 s.d.)					
hourly income	18.82	11.42	7.41	0.63	11.75
yearly income	38,939	18,911	20,028	1,205	16.62
Wage-employed in Puerto Rico vs Self-employed in U.S.					
Caliper Matching (1 s.d.)					
hourly income	18.78	11.77	7.00	0.53	13.20
yearly income	38,888	23,147	15,741	1,109	14.20
Caliper Matching (2 s.d.)					
hourly income	18.78	11.45	7.33	0.53	13.88
yearly income	38,888	22,331	16,557	1,104	14.99
Caliper Matching (0.5 s.d.)					
hourly income	18.78	11.98	6.80	0.53	12.75
yearly income	38,888	23,665	15,223	1,112	13.70
Wage-employed in U.S. vs Self-employed in U.S.					
Caliper Matching (1 s.d.)					
hourly income	18.78	16.82	1.96	0.58	3.38
yearly income	38,912	33,864	5,048	1,129	4.47
Caliper Matching (2 s.d.)					
hourly income	18.78	16.57	2.21	0.56	3.92
yearly income	38,912	33,217	5,695	1,120	5.08
Caliper Matching (0.5 s.d.)					
hourly income	18.78	17.08	1.70	0.59	2.89
yearly income	38,912	34,447	4,465	1,136	3.93

Independent variables, internal migration

Variable	$SE_{stayers}$	SE_{movers}	Variable	$SE_{stayers}$	SE_{movers}
age < 20	0.07	0.19	industries		
	(0.25)	(0.39)	agriculture, forestry, etc.	0.09	0.05
age 20-30	0.17	0.26		(0.29)	(0.22)
	(0.37)	(0.44)	mining	0.00	0.00
age 30-40	0.27	0.21		(0.05)	(0.04)
	(0.45)	(0.41)	construction	0.17	0.15
age 40-50	0.30	0.20		(0.38)	(0.35)
	(0.46)	(0.40)	manufacturing	0.05	0.04
age 50-60	0.19	0.14		(0.21)	(0.20)
	(0.40)	(0.35)	transport/communcations	0.04	0.04
gender-male	0.64	0.59		(0.21)	(0.20)
	(0.48)	(0.49)	wholesale	0.03	0.02
gender-female	0.36	0.41		(0.16)	(0.15)
	(0.48)	(0.49)	retail	0.12	0.11
education-incomplete	0.01	0.00		(0.32)	(0.31)
	(0.08)	(0.06)	finance/insurance	0.08	0.08
education-primary	0.03	0.02		(0.28)	(0.28)
	(0.17)	(0.15)	business/repair services	0.11	0.12
education-secondary	0.41	0.33		(0.31)	(0.32)
	(0.49)	(0.47)	personal services	0.09	0.09
education-tertiary	0.55	0.65		(0.28)	(0.29)
	(0.50)	(0.48)	entertainment	0.01	0.02
married-no	0.29	0.46		(0.11)	(0.14)
	(0.45)	(0.50)	public administration	0.21	0.26
married-yes	0.71	0.54		(0.41)	(0.44)
	(0.45)	(0.50)	military	0.00	0.00
family size 1	0.19	0.35		(0.00)	0.00
	(0.39)	(0.48)	occupations		
family size 2-3	0.52	0.44	manager, professional	0.33	0.42
	(0.50)	(0.50)		(0.47)	(0.49)
family size 4-5	0.24	0.16	technician, salesman, etc.	0.22	0.21
	(0.43)	(0.37)		(0.41)	(0.41)
family size 5 <	0.05	0.04	service worker	0.14	0.14
	(0.21)	(0.19)		(0.35)	(0.34)
			farmer etc.	0.09	0.05
				(0.29)	(0.21)
			precision worker, craftsman	0.14	0.11
				(0.35)	(0.31)
			operator, fabricator, laborer	0.08	0.07
				(0.27)	(0.26)
			military	0.00	0.00
				(0.00)	0.00
			Observations	916,646	4,110

Standard deviations in parentheses.

Appendix to Chapter 5

Labor market discrimination of Puerto Rican immigrants across states and comparison groups

	Puerto Rican natives	Hispanic natives	white natives	Non-Hispanic white natives
Connecticut	0%	-7%	19%	17%
Florida	-24%	-31%	-14%	-13%
Georgia	-21%	-25%	-7%	-5%
Illinois	18%	5%	36%	4%
New York	-7%	-3%	54%	54%
Ohio	39%	-11%	7%	6%
Pennsylvania	79%	0%	43%	46%
average effect	negative	negative	positive	positive
specification of discrimination measures				
occupations	-	-	-	-
earnings	yearly	yearly	yearly	yearly

Only Oaxaca-Blinder decomposition discrimination coefficients significant below the 5% level have been considered. The discrimination measure took the value of 0 otherwise. Only states for which models achieved convergence are shown.

Labor market discrimination of Puerto Rican immigrants across states and comparison groups

	Puerto Rican natives	Hispanic natives	white natives	Non-Hispanic white natives
Connecticut	0%	-2%	19%	15%
Florida	-18%	-24%	-10%	-8%
Georgia	-11%	-16%	0%	2%
Illinois	13%	4%	29%	30%
New York	-6%	-4%	40%	40%
Ohio	-6%	-30%	-33%	-34%
Pennsylvania	61%	4%	38%	39%
average effect	negative	negative	positive	positive
specification of discrimination measures				
occupations	+	+	+	+
earnings	hourly	hourly	hourly	hourly

Only Oaxaca-Blinder decomposition discrimination coefficients significant below the 5% level have been considered. The discrimination measure took the value of 0 otherwise. Only states for which models achieved convergence are shown.

Labor market discrimination of Puerto Rican immigrants across states and comparison groups

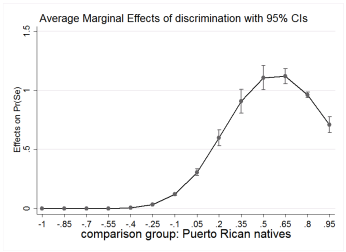
	Puerto Rican natives	Hispanic natives	white natives	Non-Hispanic white natives
Connecticut	0%	-6%	12%	7%
Florida	-20%	-27%	-14%	-13%
Georgia	-10%	0%	-2%	0%
Illinois	12%	4%	25%	25%
New York	-8%	-5%	33%	31%
Ohio	24%	-8%	-16%	-17%
Pennsylvania	60%	2%	32%	31%
average effect	negative	negative	positive	positive
specification of discrimination measures				
occupations	-	-	-	-
earnings	hourly	hourly	hourly	hourly

Only Oaxaca-Blinder decomposition discrimination coefficients significant below the 5% level have been considered. The discrimination measure took the value of 0 otherwise. Only states for which models achieved convergence are shown.

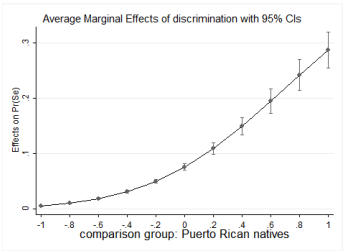
Marginal effects of various discrimination measures evaluated at value of 10% on the probability of Puerto Rican immigrant self-employment in major receiving states, by comparison groups.

	dy/dx	Std. Err.	z	P>z	[95% Conf.	Interval]	earnings	occupations
Puerto Rican natives	0.59	0.024	24.29	0.00	0.55	0.64	yearly	yes
Puerto Rican natives	0.24	0.010	23.91	0.00	0.22	0.26	yearly	no
Puerto Rican natives	0.08	0.003	23.85	0.00	0.07	0.09	hourly	yes
Puerto Rican natives	0.07	0.003	23.70	0.00	0.07	0.08	hourly	no
Hispanic natives	0.40	0.018	22.00	0.00	0.36	0.43	yearly	yes
Hispanic natives	0.33	0.013	25.02	0.00	0.31	0.36	yearly	no
Hispanic natives	0.11	0.004	24.55	0.00	0.10	0.12	hourly	yes
Hispanic natives	0.10	0.004	24.19	0.00	0.10	0.11	hourly	no
white natives	0.19	0.010	19.33	0.00	0.17	0.21	yearly	yes
white natives	0.20	0.009	22.75	0.00	0.18	0.22	yearly	no
white natives	0.06	0.003	21.87	0.00	0.05	0.06	hourly	yes
white natives	0.07	0.003	22.94	0.00	0.06	0.07	hourly	no
white non-Hispanic natives	0.19	0.010	18.69	0.00	0.17	0.21	yearly	yes
white non-Hispanic natives	0.20	0.009	21.96	0.00	0.19	0.22	yearly	no
white non-Hispanic natives	0.06	0.003	21.06	0.00	0.05	0.06	hourly	yes
white non-Hispanic natives	0.07	0.003	22.43	0.00	0.06	0.07	hourly	no

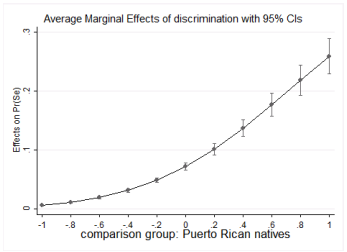
Marginal effects of various levels of discrimination on the Probability of Puerto Rican immigrant self-employment, by comparison group.



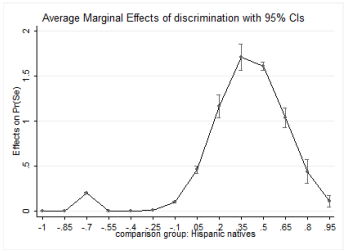
(a) yearly earnings, no occupations



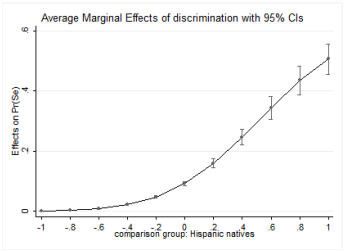
(b) hourly earnings



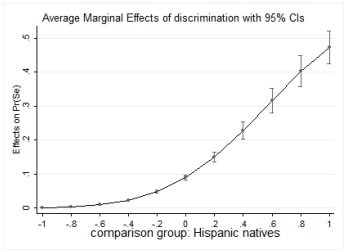
(c) hourly earnings, no occupations



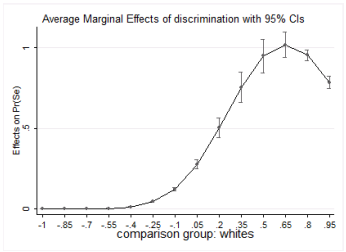
(d) yearly earnings, no occupations



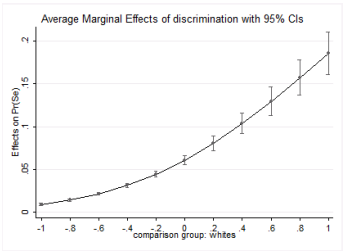
(e) hourly earnings



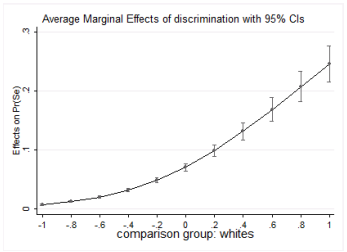
(f) hourly earnings, no occupations



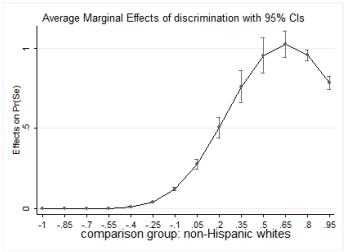
(g) yearly earnings, no occupations



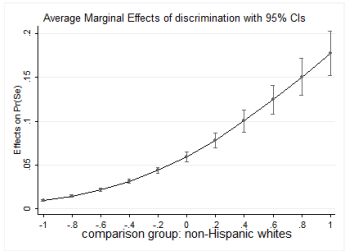
(h) hourly earnings



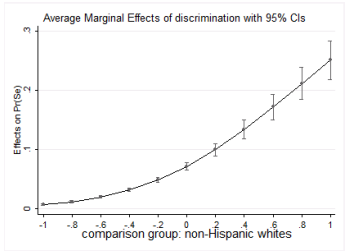
(i) hourly earnings, no occupations



(j) yearly earnings, no occupations



(k) hourly earnings



(l) hourly earnings, no occupations