ESSAYS IN MIGRATION STUDIES: REGIONAL INTEGRATION, THE GEOGRAPHY OF HUMAN CAPITAL AND ETHNIC COMMUNITIES FORMATION IN CITIES

NICOLA DANIELE CONIGLIO

Submitted in fulfilment of the degree of Doctor of Philosophy

March, 2004

Department of Economics,
University of Glasgow

ProQuest Number: 10390768

ProQuest 10390768

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ABSTRACT

The essays within this thesis investigate a series of issues of a complex and varied phenomenon: migration. In the first essay we investigate the effects of regional economic integration on the incentive for labour to spatially relocate. By introducing skill heterogeneity and localised human capital externalities in a basic, new economic geography model we provide a framework for re-interpreting the basic core-periphery model as a brain drain model. Incentives and timing of migration of highly skilled individuals depend on the process of goods market integration. When trade costs are high, as in the traditional core-periphery model, a dispersed equilibrium with no incentive to migrate is a stable outcome. The initial phase of regional economic integration is characterised by a brain drain: the highly skilled have stronger incentives to migrate and will compose the initial wave of migrants. If integration comes to a halt, a positive self-selection of the migrants (with all of the highly skilled in one region in my model) might be a stable equilibrium. Eventually, if regional economic integration proceeds, further low-skilled migrants might find it convenient to follow the same pattern. These will lead to a complete agglomeration of all footloose workers in the core region. The results of the paper show the existence of a range of trade costs for which only high-skilled workers have an incentive to migrate. Therefore the benefit of introducing labour heterogeneity in the basic core-periphery model is to enable us to explain one of the most striking features of interregional migration patterns, the positive self-selection of the migrants.

The second essay moves in the direction of learning more about the micro-level migratory behaviour of talented individuals in the context of a peripheral region. The ability of a regional system to generate human capital is crucial, but it is not a sufficient condition leading to economic success. In this work we emphasise the importance of maintaining and attracting highly skilled and educated individuals and the challenges that migration flows pose for policymaking aimed at fostering human capital accumulation in peripheral regions. We employ a unique data set generated through a postal survey designed and conducted by the author. The focus of our analysis is on the micro-level location decisions of a sample of highly educated and skilled individuals residing in Basilicata, a small Italian Mezzogiorno region, who have benefited from a locally funded human capital investment policy. Thanks to the quality and richness of the data employed, we are able to conduct a very detailed analysis of the migration propensities on the basis of individual's characteristics. We find that even within a fairly homogeneous group of highly skilled and educated individuals the probability to move is higher for the most talented. Individuals migration propensity considerably differ according to the subjects and the locations of attendance of their studies. In addition, we investigate what are the regional characteristics that make a region attractive for highly educated and talented individuals. The results of the analysis constitute a valuable knowledge that can be used in a strategic way to formulate talent attraction and retention policies.

In the last essay, we extend previous models of migration networks and ethnic communities formation by considering migration as ethnic-community-wide phenomena where established migrants strategically provide support to newcomers. The incentive to provide support is associated with positive externalities which new waves of migrants might have on migrants already settled in the host location. Culturally-based tastes for particular goods and services generate an ethnic consumer demand and only individuals from the same ethnic community have the skills or the "insider's information" required to provide these goods (protected market). If the ethnic population is large enough, an ethnic sector will emerge and eventually grow as the ethnic population expands further.

According to the degree of preferences toward ethnic consumption, the mobility costs of the source locality population and congestion costs (hostility externalities) in the host location, alternative scenarios may arise. These scenarios provide a possible explanation of why different groups of migrants show different migration dynamics.

ACKNOWLEDGEMENTS

CONTENTS

TABLES

FIGURES

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 1

Introduction and Plan of the Research

By Nicola Daniele Coniglio

Migration is a complex and varied phenomenon. Its determinants and consequences have attracted the interest of scholars in different disciplines, such as demography, economics, sociology and urban planning, just to name the most obvious. This interest is hardly surprising given the importance of migration as a powerful mechanism of change for our economies, our societies and, in general, the "space" where we live.

Subsequent waves of researchers have enlarged the field of observation and investigated many different aspects of human migration, considerably expanding our knowledge on this fascinating subject. A rather incomplete list of issues addressed by recent literature would include: the determinants of internal and international migration flows, the economic performance of migrants relative to non-migrants, the role of migration as an adjustment to regional economic imbalances, and the effects of migration on both those left behind and on natives in the place of destination.

This thesis focuses on a selected set of issues in this vast literature on migration studies. My research journey has been strongly influenced by my personal experience and by the actual and historical socio-economic environment of the area where I have spent most of my life, the so-called Italian Mezzogiorno. I will now outline in more detail the different stages of my research and describe the contents of the pages to follow.

The starting point of my research enterprise literally coincided with the moment I took the train from Bari (south-east of the Italian Mezzogiorno) to Milan (the economic "core" of Italy), namely the beginning of my travel to Glasgow where in 1999 I started my doctoral programme in economics.

It was not the first time I took the Bari-Milan line, but it was the first time that I realised how different is the flow of travellers from South to North compared to those travelling the opposite way. The stereotype of a South-to-North traveller, with the obvious simplifications, is a young woman (or man) with a university degree in her pocket. The luggage she is carrying is full of gastronomic delicacies prepared by her mother, the precious stock that will probably not last until the next weekend spent at home. She is probably on her first move to a job located in the North. The fact that you observe a laptop bag among her belonging tells you a lot on how the nature of migration has changed along this South-North route. Not many years ago, the image of a typical migrant was that of a man with a *valigia di cartone* (old cardboard suitcase), relatively unskilled and speaking almost only the local dialect. To sum up, this train slowly carries northward its load of young and highly educated individuals.

When the journey is taken in the opposite direction, from North to South, the average age dramatically increases along with a reduction of the average level of education and skills. On the train you will find tourists travelling South, to enjoy the sun and the above-mentioned gastronomic delicacies. Along with them will be returning former migrants, those of the cardboard suitcase, who plan to spend their retirement in their village of origin.

Obviously, this impressionist picture I formed is far from being a scientific observation but it was enough to stimulate my interest and curiosity on the issue of migration of educated and skilled individuals from peripheral regions.

Several studies have demonstrated that some groups of individuals are more geographically mobile and more reactive to regional wage differentials than others. In the migration literature the fact that economic migrants tend to be, on average, more skilled, young, educated and entrepreneurial than similar individuals who choose to remain in their place of origin, is a standard proposition (known as the *positive self-selection of the migrants*). A better understanding of the pattern of migration is important for assessing the economic and sociological consequences for the origin and destination regions. The more migrants are favourably selected, the more beneficial will be their impact on the destination economy. On the other hand, such a selection will have, in general, a more adverse effect on the origin region (known as the *brain drain*).

The issue of the brain drain has undergone a recent revival. The debate is fierce not only in developing countries where the brain drain has been historically severe, but also among policymakers and economists in developed countries. The Italian Mezzogiorno represents a reality where this debate is particularly intense.

In my opinion, the phenomenon of brain drain within a country goes hand in hand with the phenomenon of agglomeration of economic activities into core regions and regional centralisation of human capital. The fundamental asymmetry and selectivity between interregional migration flows in the Italian Mezzogiorno, as in other peripheral regions, are the evident consequences of regional economic disparities. Is it possible that asymmetry is not only the consequence but also one of the major causes of regional economic imbalances? The first part of my research is devoted to this issue. If human capital flows from where it is scarce to where it is abundant, as Lucas (1988) argues, migration might worsen regional

economic imbalances rather than, as a neoclassical model would predict, reduce such imbalances.

The economic geography framework offers useful insights into the analysis of the brain drain and its consequences. The strength of the economic geography explanation of the brain drain lies in its explicit general equilibrium nature which gives internal consistency to the analysis. Agglomeration of highly skilled workers in core regions is a consequence of economic integration and the reduction of trade costs in the tradable sector. Spatial unevenness and centralisation of human capital in core locations might result even when regions have identical starting points.

In my first chapter, by introducing skill heterogeneity and localised human capital externalities in a basic, new economic geography model (Krugman, 1991) I provide a framework for re-interpreting the basic core-periphery model as a brain drain model.

Incentives and "timing" of migration of highly skilled individuals depend on the process of goods market integration. When trade costs are high, as in the traditional core-periphery model, a dispersed equilibrium with no incentive to migrate is a stable outcome. The initial phase of regional economic integration is characterised by a brain drain: the highly skilled have stronger incentives to migrate and will compose the initial wave of migrants. If integration comes to a halt, a positive self-selection of the migrants (with all of the highly skilled in one region in my model) might be a stable equilibrium. Eventually, if regional economic integration proceeds, further low-skilled migrants might find it convenient to follow the same pattern. These will lead to a complete agglomeration of all footloose workers in the core region, as in the Krugman model.

The results of the paper show the existence of a range of trade costs for which only high-skilled workers have an incentive to migrate. Therefore the benefit of introducing labour heterogeneity in the basic core-periphery model is to enable us to explain one of the most

striking features of interregional migration patterns, the positive self-selection of the migrants. Another important implication of the model is the existence of a persistent wage and productivity differential between the core and peripheral regions, which is due to two main factors: the sorting of highly skilled individuals in the core and localised human capital externalities associated with the skilled density in the core. This prediction has been supported by a recent strand of empirical literature to which I refer in more detail later in this thesis.

In my opinion, the simple framework developed in this initial part of my research journey, contributes an explanation of the nature of actual interregional migration between core and peripheral regions across the European Union. It has been argued that the size of the interregional flows between and within European countries is rather limited, and that, therefore, migration represents a weak channel of change for the internal economic geography of the area. The relevance of the qualitative aspects of these flows, namely the skill composition of the flows, is often neglected. What are the skill composition and the geography of migration flows between European core and peripheral regions? Are peripheral regions experiencing a brain drain? In order to answer these important questions more research needs to be undertaken and more detailed data need to be publicly available in order to shed light on the migration behaviour of highly skilled individuals within Europe. The second step of my research journey moves in the direction of learning more about the micro-level migratory behaviour of talented individuals in the context of a peripheral region.

Human capital and migration, from theory to empirics: learning from a case study

The importance of human capital as a key source of value added, innovation and economic growth is widely acknowledged by both economists and policymakers. The relevance of

human capital in modern economies, and the divergence between its private and social return, justify public policy interventions aimed at subsidising human capital accumulation. The regional dimension is increasingly becoming central in this process of knowledge creation: local governments can directly affect individuals' decision making by subsidising human capital formation.

In the light of my initial part of the research, the scope for policy might vanish if locally formed human capital is lost through migration. The public benefit of human capital investment does not always accrue where it is funded or where policy makers intended human capital to produce its positive effects.

The ability of a regional system to generate human capital is crucial, but it is not a sufficient condition for economic success. It is of considerable importance to maintain and attract highly skilled and educated individuals and migration flows pose challenges to policymaking aimed at fostering human capital accumulation in peripheral regions. In the second part of my research, namely Chapter 2, the effort has been directed to shed light on the factors shaping the economic geography of human capital.

The focus of my analysis is on the micro-level location decisions of a sample of highly educated and skilled individuals residing in Basilicata, a small Italian Mezzogiorno region, who have benefited from a locally funded human-capital investment policy. The regional policymakers, in recognition of the importance of human capital as a key ingredient for regional growth, have given generous subsidies since the beginning of the 1990s to young graduates who want to attend a postgraduate course either inside or outside the region. The sample of individuals under analysis is far from being representative of the regional population, but the biased nature of the sample (highly skilled and educated individuals) makes the analysis a natural experiment for (i) studying the factors which affect the location decision of this important stratum of the population; and (ii) assessing the ability of the

regional system in a peripheral region not only to *generate* human capital, but also to *retain* it.

Through a postal survey, I generated a dataset containing a rich set of individual-level information. The conduct of the survey and the design of the questionnaire employed are presented in detail in Chapter 4. I believe that a significant part of the value added of my analysis is associated with the richness of the data collected; a unique tool for studying the micro-level migration decisions of highly skilled individuals in a peripheral region.

By using micro-econometric techniques, I address the following questions:

1) Who are the migrants? Are there significant differences in the personal characteristics of migrants and non-migrants?

2) What are the main factors influencing migration?

3) Where do migrants go? Is the human capital attracted toward core regions? What are the regional characteristics that make a region attractive for highly educated and talented individuals?

The existing migration literature points out the likelihood that the highly skilled and well-educated are more likely to be mobile than the rest of the population. However few studies have explored the factors underlying individual migration decisions for this group. If policy makers are to capitalise on highly skilled and educated labour, regarded as one of the most important resources in modern economies, a first step has to be to understand why they migrate. This knowledge can then be used in a strategic way to inform talent attraction and retention policies.

In my study, I am able to show that even within a quite homogeneous group of individuals, there are key differences which affect the chances of migration. In order to make a region attractive to talented individuals, strong economic performance and information on the availability of job opportunities are important pre-requisites, although they do not

necessarily guarantee success. Among the factors explaining individuals' location decisions, quality of life is very important. For this reason, policy interventions aimed at improving quality of citizens' lives should be regarded as key strategic objectives. A large degree of criminality significantly discourages the potential location decision of talented individuals and investors. Improving quality of life can lead to a substantial pay-off in terms of future growth and prosperity of peripheral regions.

Thanks to the quality of the data at my disposal, I am able to assess the importance of information and past migration experiences on an individual's migration behaviour. The probability that an individual will select a given location increases the better the information available on job opportunities in that location. Lack of information, or indeed poor quality information on local economic opportunities, represents a market failure that can be at least partly corrected by policy measures. Actions aimed at this objective might therefore usefully be considered to be strategic in trying to retain and attract human capital.

My results have strong policy implications for the local economy under analysis. However, in my opinion, they also have a general validity for regions facing similar circumstances, such as many peripheral regions across Europe.

Migration networks and ethnic communities in cities

The migration literature argues that the stock of established migrants originating from the same location is almost always found to be an important explanatory variable of subsequent migration streams (see Greenwood, 1969). This finding suggests that mobility costs are endogenous and decreasing in the stock of migrants from the same source area. In my research, I find this result to hold even for a specific stratum of the population, i.e. highly skilled and educated individuals.

Why do established migrants from the same community affect mobility costs for subsequent incoming migrants? Established migrants often provide direct assistance to new migrants. They might help finance transportation costs, provide housing or pre-arrange jobs. In addition, established migrants in the network supply information, which allows recipients to assess more carefully the expected returns from migration. In my empirical study, described above, the role of information in individuals' location decision proved to be a fundamental explanatory variable. This finding, which might seem obvious is, I believe, often discounted by economists interested in migration, and it has channelled my research interests toward a new direction. Nelson (1959) has classified information on which potential migrants form their expectations into two main categories: specific and generalised information. Specific information implies knowledge about particular opportunities, while generalised information concerns the awareness about the attractiveness of opportunities in a location. Established migrants have a fundamental role in conveying both types of information. In particular, by providing specific information and direct assistance they considerably reduce migration costs for incoming migrants.

The interplay between subsequent waves of migrants and the set of complex relations, which characterises this interplay, is in the literature often called *migration network* or *migration chain*. Taking into account the effect of migration networks implies a shift of the focus from migration as individual optimising behaviour to migration as the result of mutual interdependence of a wider number of individuals.

I have always found fascinating the way migrants during the initial part of the last century selected their destinations. Migrants from neighbouring villages in the source country, choose completely distinct destinations, different urban areas, different destination countries, indeed, often different continents. Clearly, migrants belonging to source areas characterised by common underlying "push" factors were attracted to destination areas with distinct "pull"

factors. The clearest explanation of such a complex migration pattern is the existence of community level network externalities. Migration must be seen as a community-wide phenomenon rather than the independent decision of an individual taken in isolation.

Why is this different perspective about the migration dynamics important? The existence of a migration network and its mechanisms has a profound impact on the size and geography of migration flows. For instance, while a large set of potential destinations exists, migration is often channelled into only a very limited number of them. Migration flows originating from a given source area are often extremely spatially concentrated giving rise to the various Little Italys, Little Karachis or Chinatowns which characterise urban areas around the world.

I propose a theoretical model that explains both established migrants' support for newcomers and the formation of ethnic communities on the basis of the existence of "ethnic consumption externalities" (Chapter 5). Incoming migrants from the same ethnic community have an expansionistic effect on the number of ethnic-specific non-tradable goods and services available to established immigrants in the host region. Culturally-based tastes for particular goods and services generate an ethnic consumer demand and only individuals from the same ethnic community have the skills or the "insider's information" required to provide these goods (protected market). If the ethnic population is large enough, an ethnic sector will emerge and eventually grow as the population expands further. Migrant communities in almost every big city around the world provide a wide range of ethnic goods and services. Ethnic specific goods include for instance preparation of ethnic foods, banking and financial services, ethnic media such as magazines, radio, newspapers, religious services, organised unions, political groups and usually a vast array of migrant associations.

The main questions I address are the following:

1) When will established migrants provide support to newcomers?

2) What is the resulting size of an ethnic group?

3) Is the resulting size equal to the optimal size?

4) When does an ethnic productive sector arise?

The model is able to produce a rich set of scenarios. In my opinion, the multiplicity of possible scenarios provides a plausible account of divergent migration experiences showed by different ethnic groups and in different episodes of human history.

References:

Greenwood, M. J., 1969, An Analysis of the Determinants of Geographic Labor Mobility in the United States, *Review of Economics and Statistics* 51 (2), 189-194.

Krugman, P., 1991, Increasing Returns and Economic Geography, *Journal of Political Economy* 99, 483-499.

Lucas, R. E., 1988, On the Mechanics of Economic Development, *Journal of Monetary Economics* 22, 3-42.

Nelson P. (1959), Migration, real income and information, *Journal of Regional Science*, Vol. (1), 43-74

Chapter 2

Migration of the Highly Skilled: An Economic Geography Model of the Brain Drain

Nicola D. Coniglio[*]

1. *Introduction*

Migration of highly skilled individuals has assumed an increasing importance in recent decades reflecting the ongoing process of world market integration and the considerable reduction in costs of mobility.[1] While some amount of mobility is beneficial not only to the recipient region but also to the source region if the latter is to integrate in the global economy (Saxenian et al. 2002; Rauch 1999), large outflows of skilled individuals might have a negative impact on the economic performance of the source region. This is known as the *brain drain*.

The issue of the brain drain has undergone a recent revival. The debate is fierce not only in developed countries where the brain drain has been historically severe but also among policymakers and economists in developed countries.[2] The migration of highly skilled individuals raises concerns not only when flows take place between nations but also within regions in a country. The phenomenon of a brain drain within a country goes hand in hand with the phenomenon of agglomeration of economic activities into core regions and regional centralisation of human capital.[3]

Several studies have demonstrated that some groups of individuals are more geographically mobile and more reactive to regional wage differentials. In the migration

---

[1] Although restrictive immigration laws which cyclically emerge in many developed countries constitute important barriers to international labour mobility, it is a matter of fact that the process of globalization and the exponential reduction in the cost of acquiring information on distant locations have considerably reduced migration costs. This is particularly true for highly skilled individuals since a number of developed countries have undertaken policies that facilitate and encourage admission of these workers.

[2] Examples of the relevance of the brain drain issue in developed countries are numerous. In Italy, for instance, the debate is both on the considerable flows of highly skilled and educated from the South to the North of the country and on migration of talented individuals abroad (Becker et al.2003). The issue of a brain drain from Canada to the US is well represented in the September 1999 issue of *Policy Option*.

[3] Chapter 3 of the thesis studies the migration decision of talented individuals who benefited from a human capital investment policy adopted by an Italian Mezzogiorno region. The results show a clear process of regional centralization of human capital which is a common feature in many European countries.

literature the fact that economic migrants[4] tend to be, on average, the more skilled, young, educated and entrepreneurial than similar individuals who choose to remain in their place of origin, is a standard proposition generally referred to as *positive self-selection of the migrants*.[5] A better understanding of the pattern of migration is important for assessing the economic and sociological consequences for the origin and destination regions. The more migrants are favourably selected, the more beneficial will be their impact on the destination economy. On the contrary, the more positively selected are the migrants the greater, in general, will be the adverse effect in the origin region. Although the adverse effect of the brain drain on the source region is a common belief in the literature, we acknowledge that a recent strand of research maintains a more optimistic view in this regard. In a separate section of this work we will briefly present and discuss the "beneficial brain drain" hypothesis.

The economic geography framework offers useful insights to the analysis of the brain drain and its consequences. As pointed out by Commander et al. (2004) the strength of the economic geography explanation of the brain drain lies in its explicit general equilibrium nature which gives internal consistency to the analysis. Agglomeration of highly skilled workers in core regions is a consequence of economic integration and reduction of trade costs in the tradable sector. Spatial unevenness and centralisation of human capital in core locations might result even when regions have identical starting points.

In the basic model of Krugman (1991), the interaction between interregional labour migration, scale economies, transport costs and a spatially immobile source of demand generates agglomeration and dispersion forces. Since some factors of production are spatially mobile, when transport costs fall below a critical level, agglomerative forces become strong

---

[4] We refer to economic migrants as individuals who move from one place of work and residence to another, both within or across countries, on the basis of a decision taken comparing their own economic opportunities in the origin and destination locations. In this way we want to distinguish those migrants from asylum seekers and those migrating because of others decisions (such as family reunification).

enough to give rise to a core-periphery structure. Therefore, even if regions are a priori identical, they can endogenously differentiate into an industrial core and a de-industrialised periphery due to migration-induced demand linkages. The standard economic geography model cannot be directly applied to analyse the brain drain issue since the assumed homogeneity of the footloose factors of production implies that we inevitably ignore, by definition, the "brain" component (Commander et al. 2004).

In this paper, by introducing skill heterogeneity and localised human capital externalities into the Krugman (1991) model we are able to re-interpret the basic core-periphery model as a brain drain model. Two types of manufacturing workers exist in our simplified economy: low and high-skilled, where skills are associated with the efficiency unit of labour offered in the labour market. We assume that a low-skilled worker is able to supply one unit of labour effort. High-skilled workers are potentially more productive than the unskilled ones. Interaction between high-skilled manufacturing workers in a region increases the productivity of each worker by means of a knowledge diffusion process (we assume that geographical proximity is needed). In the model the skill premium is endogenous and increasing in the regional share of high-skilled workers.[6] Migration of the highly skilled is at the same time cause and consequence of the endogenous regional real wage differential.

The regional level of skills can be interpreted as the stock of human capital in the economy. According to Lucas (1988) the accumulation of human capital is the source of positive spillovers. It is reasonable to argue that workers' skills are augmented through learning and exchange of ideas, and that workers, therefore, increase their productivity by

---

[5] For a review of theoretical and empirical contributions on this topic see Chiswick (2000). Borjas, Bronars and Trejo (1992) find evidence of self-selection studying internal migration in the U.S. In fact migrants with higher educational levels appear to be attracted toward regions with higher returns to education.

[6] According to Glaeser (1998), firms choose to locate in cities and pay the higher wages and suffer congestion costs because workers in cities are more productive: "...if workers weren't more productive firms would leave cities altogether and hire elsewhere. Since the urban wage premium appears to be a centuries-old phenomenon, we must assume that over the long run, firms are quite willing to pay these higher wages".

interacting with those around them. As a result, migration of a worker from a region where the average level of human capital is low to one where the average level is high will raise his productivity. Geographic proximity is crucial since it allows ideas to travel more rapidly; the impact of localised externalities weakens with distance. As proposed by Kremer (1993), many production processes consist of multiple tasks, all of which must be successfully completed for the product to have a full value. The matching of skilled workers in the production process increases the probability of successfully performing those tasks. In equilibrium skilled workers are clustered in the core region. This result is consistent with a series of stylised facts, the most significant being, the substantial wage and productivity differences between rich and poor regions.

Incentives for migration of highly skilled individuals and "its timing" depend on the process of goods market integration. When trade costs are high, a dispersed equilibrium with no incentive to migrate is a stable outcome, a result in common with the traditional core-periphery model. The initial phase of regional economic integration is characterised by a brain drain: highly skilled have stronger incentives to migrate and will compose the initial wave of migrants. If integration comes to a halt, a positive self-selection of the migrants (with all highly skilled in one region in our model) might be a stable equilibrium. Eventually, if regional economic integration proceeds further, low-skilled migrants might find it convenient to follow the same pattern. These will lead to a complete agglomeration of all footloose workers in the core region as in the Krugman model.

The results of the paper show the existence of a range of trade costs for which only high-skilled workers have an incentive to migrate. Therefore the benefit of introducing labour heterogeneity into the basic core-periphery model is to enable us to explain one of the most striking features of interregional migration patterns, the positive self-selection of the

migrants. Another important implication of the model is the existence of a persistent wage and productivity differential between the core and peripheral regions, which is due to two main factors: sorting of highly skilled individuals in the core and localised human capital externalities associated with skilled density in the core. This prediction has been supported by a recent strand of empirical literature.

A recent work by Giannetti (2003) addresses the issue of positive self-selection of migrants and regional convergence using a different framework from the one used here. In an overlapping-generations model of a two-regions economy characterised by localised human capital externalities and skills heterogeneity, the author shows that an initial economic disadvantage (lower capital abundance) might give rise to a poverty trap since self-selection of migrants causes regional differences in total factor productivity and hampers convergence. Cost-of-living differentials, arising as a consequence of the Balassa-Samuelson effect, explain migration flows of low-skilled workers from the high-income to the low-income region. While we work on very related issues, our perspective is different since we want to investigate the effects of a process of regional economic integration on the spatial distribution of skills and regional convergence. In addition, as in almost all new economic geography papers, our point of departure is represented by two initially identical regions. No underlying differences are assumed.

Another related contribution is that of Mori and Turrini (2004). As in our contribution they use a new economic geography framework to study spatial clustering of individuals with heterogeneous skills. The mechanism and the main assumptions of their model greatly differ from our approach. In their modification of the Krugman (1991) core-periphery model, many heterogeneous worker-sellers offer goods which are both horizontally (variety) and vertically (quality) differentiated. The quality of a variety depends on the skill level of the workers who produce it. Free entry ensures that skilled workers reap all rents associated with their skills. A

crucial assumption in their model is the existence of a fixed communication cost together with the traditional (proportional) iceberg trade costs. Communication costs imply that imported goods' perceived quality is discounted by a fixed amount, irrespective of their true quality. In turn, skilled agents are proportionally less affected by distance compared with the less skilled. These non-convexies in the technology for selling goods to distant locations imply that the more skilled are relatively less dependent upon local markets. They have an incentive to move toward locations where other skilled workers reside since a wider range of high quality goods will be available (lower price index). Conversely, less skilled workers tend to cluster in a region where their local sales are abundant but where, at the same time, locally provided goods are of relatively low quality. Their model predicts an inevitable regional inequality since spatial sorting/skill segmentation is always a feature of stable equilibria. We base our model on a different premise, i.e. the existence of human capital externalities, and specifically investigate the effect of regional economic integration on spatial equilibria. Regional inequality and skill segmentation is not inevitable nor automatic in our model but depends on the degree of integration (proxied by the size of transport costs).

The paper is organized as follows. In section 2 we present recent empirical evidence on agglomeration and human capital externalities. Section 3 develops the model. In section 4 we examine the possible spatial equilibria and their features. In section 5, we briefly discuss a recent strand of research which advocates the idea of a beneficial brain drain. The last section concludes.

2. Agglomeration and Human Capital Externalities: empirical evidence

In our model, the interaction between high-skilled workers in a region increases the productivity of each worker by means of a knowledge diffusion process. Regions with a

higher average level of human capital are therefore more attractive for high-skilled workers. The skill premium is endogenous and increasing in the regional share of high-skilled workers. In our model, therefore, migration flows of high-skilled individuals and the extent of human capital externalities are closely interrelated phenomena. One of the implications of the model is the existence of a wage and productivity differential between the core and peripheral regions.

Economic theory gives several explanations on why wages may be different across regions in the short run, one of the main reasons being the existence of asymmetrical regional shocks and, over time, divergences of regional business cycles. In the long run, after the necessary adjustments in local labour markets have taken place, regional wages should, in principle, converge.

One possible explanation for the observed lack of convergence in the long run is linked to the existence of differences in the regional "endowment" of amenities. Local amenities, such as good climate, favorable physical morphology of the area, air and water quality, may affect regional wages through two main channels. First, through a positive (or negative) effect on consumer utility. For example if consumers value the warm and sunny weather of Southern Italy, they will require a wage premium to live and work in rainy and cloudy climate of Northern Italy. As a result, *ceteris paribus*, we expect workers of similar characteristics to be willing to accept lower salaries in regions with pleasant weather conditions. Secondly, local non-exclusive amenities may also have a direct influence on labour productivity. If the effect is positive we will expect higher wages. The attractiveness of the area, due to a positive regional wage differential, will induce migration, and therefore these productivity differences will be capitalised in land rents.

The impact of exogenous regional characteristics on the spatial variation in factor prices is generally evaluated using hedonic price estimation techniques (Roback 1982). As

Hanson (2001) points out, although empirical evidence supports the hypothesis that local amenities contribute to explaining regional differences of factor prices, some evidence shows that considerable differences remain even within regions with a similar endowment of exogenous amenities.

Some theories, based on the existence of localized human capital externalities, argue that differentials in factor remuneration may be persistent over time if regions have different levels of human capital (see Eaton and Eckstein 1997). The main idea is that the level of skills, education and experiences of the local labour force positively affect workers' productivity. Interaction between skilled individual enables flows of ideas, diffusion of best practices and, in general, the possibility to benefit from a local knowledge stock. As a consequence, high-skilled workers from other areas will also be attracted toward regions with higher returns to skills. Therefore, a positive self-selection of migrants (see Borjas et al. 1992) may generate a self-sustaining mechanism of growth. According to the theory the level of local human-capital has an important role in explaining spatial differences in wages and housing prices. Recent empirical studies have supported this prediction using micro data. Rauch (1993) employs US Census data on wages and human capital of individuals in 237 cities in 1980 to estimate externalities in cities using individual wage-regressions. The notion of human capital defined by Rauch contains both education and work experience components. His results suggest that the external effect of a one-year increase in average schooling in cities has a positive and statistically significant effect on wages of workers in the same city of around 4%. These results hold even when he controls for the effects of other factors such as R&D investment policies that favour some cities, as well as university concentration in urban areas. More recently, Ciccone and Peri (2000) find that a one-year increase in average schooling in US cities raises aggregate productivity by 8 to 11%.[7]

---

[7] For a survey on empirical work on human capital externalities in cities see Moretti (2004).

How are externalities transmitted across regions? Is there any evidence that their effects are weaker over larger distances? Ciccone and Hall (1996) use data at the county level to see if variations in population density can explain the large discrepancies in productivity levels across US.[8] In their analysis, they control for the endogeneity of employment density to ensure that the correlation between density and productivity that they find is not merely the result of productive regions growing faster than less productive ones. They find that doubling employment density increases labour productivity by 6%. Their results hold even when other factors are taken into account (such as the level of public capital, the level of education and the influence of market size). According to their analysis, closer interaction between workers in a geographical unit does have a positive effect on productivity. The elasticity of wage with respect to the local employment density in studies similar to Ciccone and Hall (1996) ranges between 4 and 8% (for a recent comprehensive survey see Rosenthal and Strange, 2004). These estimates of agglomeration economies might be considerably upward biased according to a study by Combes et al. (2004). A simple but important reason for spatial differences in wages is due to spatial differences in the composition of the workforce. Sorting of workers by skills across areas implies a different average wage. Combes et al. (2004) perform an empirical analysis on the determinants of wage disparities across French employment areas employing a large panel of French workers comprising 20 million observations between 1976 and 1998. The authors are able to take into account, under the same framework, of the effects on wage disparities of: (i) geographical skill sorting; (ii) spatial differences in endowments; (iii) agglomeration externalities. Spatial sorting of skills accounts for most of the explained disparities in wages. Their results highlights the considerable importance of a positive self-selection in migration flows in explaining persistence in wage and income regional inequality. It is interesting to note that their findings do confirm the importance of

agglomeration economies in densely populated areas.[9] Skilled workers tend to cluster together in larger and denser locations; the skill composition effect magnifies the agglomeration effect on spatial wage disparity.

Indirect evidence supports the argument that the relevance of spatial influence decays as distance increases. In fact, the geographical distribution of income and unemployment both in absolute values and variations over time is strongly correlated across neighbouring regions (see Overman and Puga, 1999, and Quah, 1996).

3. The Model

We consider a world economy consisting of two regions (1,2). There are two sectors: agriculture and manufacturing. Agriculture is perfectly competitive and produces a homogeneous good employing interregionally immobile farm labourers (sector-specific factor). In our "brain-drain" model this reflects the fact that some individuals in the population are perfectly immobile.

Manufacturing is a monopolistically competitive sector producing a variety of differentiated products with increasing-return-to-scale technology and employs interregionally mobile workers. We assume the existence of two types of workers in the manufacturing sector, which differ only in terms of the level of efficiency unit of labour supplied. In what follows we define in details the building blocks of our model.

3.1 Consumers' behaviour

All individuals share the same Cobb-Douglas utility function:

---

[8] According to the authors, in the U.S. a worker in the most productive state is two-thirds more productive than a worker in the least productive state.

$$U = M^{\mu} A^{1-\mu} \qquad (1)$$

where M is a quantity index of consumption of manufactured goods and A is consumption of the agricultural good. Therefore μ is the expenditure share of manufactured goods. The manufacturing aggregate M, is a sub-utility function of a discrete number of varieties defined by a constant elasticity of substitution function:

$$M = \left[\sum_{1}^{n} m_i^{\rho} \right]^{1/\rho} = \left[\sum_{1}^{n} m_i^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \qquad (2)$$

where n is the number of varieties produced, m_i is consumption of each variety and ρ is a parameter representing the intensity of the "love for variety" in the manufacturing sector. The constant elasticity of substitution between any two varieties is $\sigma = 1/(1 - \rho)$, $(\sigma > 1)$.[10]

The consumers maximize (1) subject to the following budget constraint:

$$P_A A + \sum_{1}^{n} m_i p_i = Y$$

where Y is income and P_A, p_i are respectively, the prices of the agricultural product and prices for each variety of the manufacturing aggregate.

A two-stage budgeting procedure can be applied. The first step in the consumer's problem is to choose each m_i in order to minimise the cost of attaining a given M:

---

[9] Their estimated elasticity of wages with respect to local employment density (which is around 2%) is considerably lower than those found in similar studies (from 4 to 8%).

[10] Then number of varieties produced, n, is assumed to be large. The use of integer is formally correct and qualitatively equivalent to the use of a concept of a continuum of goods. Later, in Chapter 5, it will be more convenient to treat n as a continuous variable.

$$\min \sum_1^n p_i m_i \quad \text{s.t.} \quad M = \left[\sum_1^n m_i^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

Solving the first-order conditions yields:

$$m_i = \left(\frac{p_i}{p_j} \right)^{-\sigma} m_j$$

and by substitution of this into the budget constraint we obtain:

$$M = \left[\sum_1^n \left(\frac{p_i}{p_j} \right)^{1-\sigma} m_j^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \quad \text{which implies:}$$

$$m_j = \left[\frac{p_j}{P} \right]^{-\sigma} M \tag{3}$$

Equation (3) is the compensated demand function for the jth variety, where P is the *manufactured goods price index*, a measure of the minimum cost of purchasing a unit of the composite index M:

$$P = \left[\sum_1^n p_i^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{4}$$

The second step of the consumer's problem is to choose the optimal allocation of income between A and M so that the utility is maximised. Maximising $U = M^\mu A^{1-\mu}$ subject to $PM + P_A A = Y$, gives us the uncompensated demand for A and for each variety, m_j:

$$A = (1-\mu)Y / P_A \tag{5}$$

$$m_j = \frac{p_j^{-\sigma}}{P^{-(\sigma-1)}} Y\mu \tag{6}$$

From the consumer's utility maximisation problem we can also express the indirect utility function, substituting (5) and (6) into (1) yields:

$$U = \frac{\mu^\mu (1-\mu)^{1-\mu} Y}{P^\mu P_A^{(1-\mu)}} \tag{7}$$

the term $P^\mu P_A^{(1-\mu)}$ can be interpreted as the regional cost-of-living index in the economy.

What is the welfare effect of an increase in the number of varieties? Assuming that all varieties are available at the same price $p_i = p_j, \forall j \in [1,...,n]$, we can rewrite the manufacturing goods price index as:

$$P = \left[\sum_1^n p_i^{1-\sigma}\right]^{\frac{1}{1-\sigma}} = p_i n^{\frac{1}{1-\sigma}}$$

Therefore, an increase in the number of available varieties reduces the manufacturing price index. Consequently, given that the indirect utility function (7), is decreasing in the manufacturing price index, the regional welfare increases as the number of varieties rises. The elasticity of substitution between varieties, σ, determines the responsiveness of the price index to a change in the available number of varieties.

3.2 Labour supply and human capital externalities

The world farmer's population employed in the agriculture sector is $1 - \mu$. Since farmers cannot regionally relocate, the supply in each region is $(1 - \mu)/2$.

The manufacturing sector employs two types of workers (i) *low-skilled workers*, and (ii) *high-skilled workers*.[11] Both types of workers are endowed with one unit of labour, and may move between regions. We assume that the global population of workers of each type is normalised to 1. The total population of manufacturing workers therefore sums to 2.

The existence of localised human capital externalities implies that the interaction between high-skilled workers in a region increases the productivity of each skilled worker by a knowledge diffusion process. The productivity of high-skilled workers, i.e. the efficiency units of labour supplied, depends upon the number of workers with similar characteristics in the regional labour force. In the model the skill premium s_i is endogenous, and may differ between regions according to the size of the regional highly skilled population $s_i = f(S_i)$, where S_i represents the regional high-skilled workers population. We formalise the positive interaction between high-skilled workers in a region as:

$$s_i = S_i^{\lambda} \tag{8}$$

---

[11] In this model, labour provided by low and high skilled is qualitatively homogeneous. Manufacturing workers of the two types are perfect substitutes.

where the parameter λ measures the strength of human capital externalities. Such externalities are increasing in the quota of high-skilled workers but at a decreasing rate. Given $\partial's_i/\partial S_i = \lambda(S_i)^{\lambda-1} > 0$, and $\partial''s_i/\partial S_i = \lambda(\lambda-1)(S_i)^{\lambda-2} < 0$, it follows that $0 < \lambda < 1$. The fact that the high skilled are more efficient in providing unit of labour is reflected in the following relation between low/high skilled competitive wages in each location:

$$w_i^s = w_i \left(1 + s_i\right) \tag{9}$$

where w_i^s and w_i, represent respectively the nominal wages of high-skilled and low-skilled workers in region i.

It is useful to distinguish between the world population of manufacturing workers and the world total supply of effective units of labour. The two measures are not identical since the geographical distribution of skilled workers (the human capital level in each region) affects the total number of units of labour supplied. Total units of effective labour in one region are given by:

$$L_i = U_i + (S_i)^{1+\lambda} \tag{10}$$

where U_i and S_i represent respectively the share of total low and high-skilled workers population, which live in region i.

It is important to note that the number of skilled workers, as a fraction of the total manufacturing labour force, is exogenously determined.[12]

Agriculture is a constant-return-to-scale, perfectly competitive sector. We normalise the unitary labour input requirement to 1. Agricultural products are freely traded and therefore agricultural prices and wages are equalised across regions and shall be our numeraire. Consequently, $P_A = w_A = 1$ in both regions.

The production of any variety of the manufactured good involves a fixed cost and a constant marginal cost:

$$l_i = \alpha + \beta x_i \qquad\qquad (11)$$

where x_i is the quantity produced and l_i the labour requirement for its production. Because of economies of scale at firm level and consumers' preferences for variety, each firm produces a differentiated product facing an elasticity of demand equal to σ.

Imported manufacturing varieties incur Samuelson's iceberg trade costs. If a variety is shipped from one region to the other, part of each unit is assumed to "melt away" during the transport, therefore only a fraction $1/\tau$ arrives at destination.[13] The introduction of this kind of trade costs implies that, if a variety is produced in region 1, consumers in the two different regions have to pay different prices. A manufacturing good produced in region 1 costs the F.O.B. price p_i for a home consumer, while the consumers in region 2 pay the C.I.F. (τp_i).[14] As a consequence, according to the number of varieties produced in each region, the manufacturing price index P may assume different values between regions:

---

[12] In subsequent research, we consider the worker's human capital investment decision and the role of public policy.

[13] $\tau \geq 1$ represents the amount of the good dispatched per unit received at destination.

$$P_1 = \left[n_1 p_1^{1-\sigma} + n_2 (p_2 \tau)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}$$

$$P_2 = \left[n_1 (p_1 \tau)^{1-\sigma} + n_2 p_2^{1-\sigma} \right]^{\frac{1}{1-\sigma}}$$

(12)

3.4 Size of the regional manufacturing industry

Production technology is the same for each variety in both locations. Manufacturing labour is the only input. We continue to assume perfect substitutability in the production function between low-skilled and high-skilled labour inputs. A firm producing a specific variety at region i, faces a wage rate w_i for each effective unit of labour, which consequently represents the low-skilled nominal wage. [15]

Given the following profit equation:

$$\pi_i = p_i x_i - \left(\alpha + \beta x_i \right) w_i$$

(13)

the profit-maximising price of a firm located in region i is a constant mark-up over the regional wage rate:

$$p_i = \left(\frac{\sigma}{\sigma - 1} \right) \beta w_i$$

(14)

Relative prices of varieties produced in regions 1 and 2 must then be:

$$\frac{p_1}{p_2} = \frac{w_1}{w_2} \qquad (15)$$

Free entry drives profit to zero, and the optimal output level is the same for each firm in any region:

$$x_i^* = \frac{\alpha(\sigma - 1)}{\beta} \qquad (16)$$

the associated equilibrium labour input is also constant and given by substituting (16) into (11):

$$l_i^* = \alpha\sigma \qquad (17)$$

Full employment of the labour force allows us to determine the number of manufacturing varieties in equilibrium[16]:

$$n_i = \frac{L_i}{l_i^*} = \frac{L_i}{\alpha\sigma} \qquad (18)$$

An increase in the regional size works through changes in the variety of goods (firms) available, leaving unaltered both the mark-up over marginal cost and the scale of individual production. The equilibrium number of firms is therefore proportional to the region's

---

[16] Since each variety is produced in one location by a single firm, n_i represents both the number of varieties and the number of firms in region i.

effective population of workers. The relative number of varieties produced in the two regions therefore reflects their relative shares of effective labour.

$$\frac{n_1}{n_2} = \frac{L_1}{L_2} \tag{19}$$

3.5 Equilibrium conditions in goods and labour markets

In equilibrium the optimal output for any firm should be equal to the demand for its product in both regions:

$$
\begin{aligned}
x_1^* &= \mu \left[\left(\frac{p_1^{-\sigma}}{P_1^{-(\sigma-1)}} Y_1 \right) + \left(\frac{(p_1 \tau)^{-\sigma}}{P_2^{-(\sigma-1)}} Y_2 \right) \right] \\
x_2^* &= \mu \left[\left(\frac{(p_2 \tau)^{-\sigma}}{P_1^{-(\sigma-1)}} Y_1 \right) + \left(\frac{p_2^{-\sigma}}{P_2^{-(\sigma-1)}} Y_2 \right) \right]
\end{aligned}
\tag{20}
$$

After some manipulation, using the pricing rule equation (14), we obtain the following expression for the nominal wages at which each firm breaks even:

$$
\begin{aligned}
w_1 &= \left(\frac{\sigma-1}{\alpha\sigma} \right) \left[\frac{\mu}{x^*} \left(Y_1 P_1^{\sigma-1} + Y_2 P_2^{\sigma-1} \tau^{1-\sigma} \right) \right]^{\frac{1}{\sigma}} \\
w_2 &= \left(\frac{\sigma-1}{\alpha\sigma} \right) \left[\frac{\mu}{x^*} \left(Y_1 P_1^{\sigma-1} \tau^{1-\sigma} + Y_2 P_2^{\sigma-1} \right) \right]^{\frac{1}{\sigma}}
\end{aligned}
\tag{21}
$$

We choose the following normalisation[17] in order to simplify the wage equation and the manufacturing price index:

$$\beta = \frac{\sigma - 1}{\sigma}$$
$$\alpha = \frac{1}{\sigma}$$

(22)

From the pricing rule equation (14) we get: $p_i = w_i$ and from equation (18) $n_i = L_i$. It is possible to rewrite the price index as:

$$P_1 = \left[L_1 w_1^{1-\sigma} + L_2 w_2^{1-\sigma} \phi \right]^{\frac{1}{1-\sigma}}$$
$$P_2 = \left[L_1 w_2^{1-\sigma} \phi + L_2 w_2^{1-\sigma} \right]^{\frac{1}{1-\sigma}}$$

(23)

and the wage equation as:

$$w_1 = (\sigma - 1) \mu^{\frac{1}{\sigma}} \left[Y_1 P_1^{\sigma-1} + Y_2 P_2^{\sigma-1} \phi \right]^{\frac{1}{\sigma}}$$
$$w_2 = (\sigma - 1) \mu^{\frac{1}{\sigma}} \left[Y_1 P_1^{\sigma-1} \phi + Y_2 P_2^{\sigma-1} \right]^{\frac{1}{\sigma}}$$

(24)

where $\phi = \tau^{1-\sigma} = 1/\tau^{\sigma-1}$ (remember $\sigma > 1$). The parameter ϕ which is a function of the trade costs can be interpreted as a parameter reflecting freeness of trade, ranging between zero for very high trade costs when $\tau \to \infty$ (autarky), and one in the case of no trade costs, i.e. $\tau = 1$.

---

[17] The scope of these normalisations, widely used in this literature, is to shift the analysis to the number of manufacturing workers and their wages in each region in order to study the equilibria in the model and its stability properties.

The price index equations (23) as noted before have an important property. The regional manufacturing price index will tend to be lower, the higher is the share of effective manufacturing labour in the region, since more varieties are produced locally without incurring therefore any transport costs (*price-index effect*).

The total regional income Y_j is given by the sum of all farmers and workers wages:

$$Y_1 = w_1 L_1 + (1-\mu)/2$$
$$Y_2 = w_2 L_2 + (1-\mu)/2$$

(25)

where $1-\mu$ is the world population of immobile farmers.

Real wages are obtained by deflating nominal wages by the regional cost-of-living index P_j^μ.[18] We assume that workers are paid linearly according to their level of skill s_i, consequently the real wages, (ω_i^u, ω_i^s respectively for low- and high-skilled workers) are:

$$\omega_1^u = \frac{w_1}{P_1^\mu}$$

$$\omega_2^u = \frac{w_2}{P_2^\mu}$$

(26)

$$\omega_1^s = \frac{w_1 [1+(S_1)^\lambda]}{P_1^\mu}$$

$$\omega_2^s = \frac{w_2 [1+(S_2)^\lambda]}{P_2^\mu}$$

(27)

The economy is assumed to reach instantaneously a short-run equilibrium for any given allocation of workers between regions. The solution of the set of equations (23)-(25)

---

[18] As the agricultural product is the numeraire.

determines w_1 and w_2 for which: (i) consumers maximise utility; (ii) profits are both maximised and driven to zero by free entry; and (iii) all markets clear.

Having defined the structure and main building blocks of the model, we now move to an analysis of the migration behaviour and the resulting spatial structure of the economy.

4. Migration Flows and Spatial Equilibria

In the long run, in the absence of migration costs, real wage differentials are the only determinant in the decision of low and high-skilled workers to move from one region to the other. The regional share of both types of workers adjusts according to the real wage difference:

$$
\begin{aligned}
\dot{U}_l &= \chi(\omega_1^u - \omega_2^u) = \chi(\omega^u) \\
\dot{S}_l &= \chi(\omega_1^s - \omega_2^s) = \chi(\omega^s) \\
\chi' &> 0
\end{aligned}
\tag{28}
$$

where χ is a function increasing in ω^u and ω^s, which represent the regional real wage differentials for low- and high-skilled workers. Regional migration flows to region 1 are positive if workers enjoy a higher level of utility (i.e. higher real wages) by moving in this region.

We would like to determine when the long run equilibrium will exhibit regional convergence (yielding a symmetric equilibrium), and when it will lead to a core-periphery structure with all manufacturing workers and the manufacturing sector concentrated in one region. In addition, we are interested in determining whether in the long run the migration pattern will exhibit the feature of a positive self-selection of the migrants. To answer these questions we have to consider the local stability of these equilibria. Let us consider the case

where the world economy is at the symmetric equilibrium. It is easy to see that if labour allocation between regions is identical, that is $S_1 = S_2 = U_1 = U_2 = 1/2$, the nominal wages are the same in the two regions, $w_1 = w_2$. The same equality holds for real wages. Assume now that as a consequence of a shock to the symmetric equilibrium $L_1 > L_2$, that is some workers are allocated from region 2 to region 1. If this change in stocks positively affects real wage in region 1 relative to region 2, it will encourage migration, as the initial symmetric equilibrium is unstable. The opposite is true if the real wage differential becomes negative.

The initial movement of a single worker has four effects. [19]

The first is the *price-index effect*, working in favour of divergence. As mentioned above, the cost of living will be lower in the country with the larger manufacturing sector, because a consumer who wants to purchase a certain bundle of manufactured goods will incur in less trade costs.

A second agglomerative force, *the home-market effect*, reinforces this effect. From (24), nominal wages in a region will tend to be higher if income in the region is high. The reason is that firms can afford to pay higher wages if they have good access to a larger market. As a consequence, the large market retains a more than proportional share of manufacturing sector.[20]

Third is the *competition effect*. The presence of more firms in the local market will increase competition to serve the regionally immobile consumers, thus tending to reduce local profits and consequently equilibrium wages. This negative effect on regional wages encourages the stability of the symmetric equilibrium.

---

[19] This adjustment mechanism assumes that entry and exit of firms occurs infinitely faster than migration. Firms are therefore always in equilibrium.

[20] The idea is not a new one but finds its root in the literature on market potential (see Harris, 1954).

Finally, the migration of a skilled worker has a positive effect on the productivity of skilled workers in the host region, and therefore a positive effect on their nominal wages. We call this last effect which is peculiar to high-skilled workers migration, *skill premium effect*.

The stability of the symmetric equilibrium is given by the relative strength of these agglomerative and dispersive forces at work. For sufficiently high trade costs, imported manufactured varieties are so expensive that it is profitable to have a symmetric equilibrium. Conversely, at low trade costs the symmetric equilibrium is always unstable. Regional economic integration has no impact on workers' incentive to re-locate until a critical level of freeness of trade is reached. When this threshold of freeness of trade, the ϕ-break, is reached the symmetric outcome becomes unstable since workers have an incentive to migrate. The *skill-premium effect* is peculiar to high-skilled workers, therefore the incentive to migrate and the relative ϕ-break will be higher for this type of workers.

Formally a stable equilibrium is any point where:

(i)   the regional wage differentials are zero and $\partial\omega^u/\partial U_i, \partial\omega^u/\partial S_i, \partial\omega^s/\partial S_i, \partial\omega^s/\partial U_i$ are all strictly negative. This is the case of a stable symmetric equilibrium where $S_i = U_i = 1/2; \forall i \in \{1,2\}$. A positive increase in the share of low or high-skilled workers in the region, negatively affects the real wage differential. Agglomerative forces are dominated by dispersive forces. Therefore, as a consequence, a counter-migration of manufacturing workers will re-equilibrate the size of the regional manufacturing labour force;

(ii) $\omega^u > 0, \omega^s > 0$ and $S_i = U_i = 1$ (or vice versa $\omega^u < 0, \omega^s < 0$ and $S_i = U_i = 0$), in the case of a core-periphery equilibrium. The entire population of low and high-skilled workers will be concentrated in the core. In this case we have the condition that in the core region $\partial\omega^u/\partial U_i, \partial\omega^u/\partial S_i, \partial\omega^s/\partial S_i, \partial\omega^s/\partial U_i$ are all strictly positive;

(iii) $\omega^u = 0, \omega^s > 0$ with $S_i = 1$ and $U_i \neq 1$ when a positive self-selection equilibrium is a

stable outcome. In this case high-skilled workers will be concentrated in the core region,

while an equilibrium without complete concentration is still stable for the unskilled workers

($\partial \omega^u / \partial U_i, \partial \omega^u / \partial S_i$ should both be strictly negative).

In **figures 1 and 2**[21] we plot, respectively, the low and high-skilled workers' real

wages differentials at different levels of trade costs.[22] This allows us to analyse how the

integration process, measured by the level of trade costs, will affect the equilibrium size of

the manufacturing sector and the average level of human capital in the regions. We report the

real wage differential for low and high-skilled workers as a function of the share of only one

type of manufacturing labour keeping the other at the symmetric equilibrium. In this way we

assess, for each type of worker, how the real wage differential reacts as a consequence of

variations in the populations of both types of manufacturing worker. Specifically, diagrams

(a) [(b)] of both figures report the real wage differentials as a function of the share of low-

skilled [high-skilled] workers in region 1, given that high-skilled [low-skilled] workers are

equally distributed between the two regions, $S = 1/2$ [$U = 1/2$].

For high trade costs ($\phi = 0.05$, or $\tau = 2.714$) the symmetric equilibrium is stable since

the cost of supplying a market by exporting is too large. Both low and high-skilled have no

incentive to migrate in region 2. When a worker migrates (either high or low-skilled), the

host region becomes less attractive than the source one. In the long run the economy

converges to a symmetric equilibrium in which manufacturing is equally divided.

---

[21] The model cannot be solved analytically. In the paper, all the figures presented are derived by numerical simulations. The values of the parameters are similar to those used in related papers. We let $\mu = 0.3$, while $\sigma = 4$; mark-up estimates are normally between 20-30%, which correspond to an elasticity of substitution σ between 6 and 4. Agglomerative externalities are chosen on the basis of the empirical evidence mentioned in section 2 of this chapter, $\lambda = 0.1$. Analytical expressions for ϕ^u , ϕ^s *break* and ϕ *self-selection*, have been derived using a procedure first introduced by Puga (1999). Maple files containing the simulation procedure and the stability analysis are available on request.

[22] Since the two regions are perfectly symmetric we restrict the following analysis to region 1.

As the economy becomes slightly more integrated ($\phi = 0.07$, or $\tau = 2.487$), the symmetric equilibrium is no longer stable for both types of workers. Consider the low-skilled first. From figure 1 it is evident that an increase in the share of both low and high-skilled workers has a negative effect on the real wage differential (the schedule is downward-sloping for both U and S). The competition effect is in this case stronger than the effect of the other two agglomerative forces. For the high-skilled worker the situation is different (figure 2). If the skilled manufacturing labour force concentrates in one region, these workers become more productive and the competition effect is more than compensated by the skill premium effect. However, migration of the low-skilled has a negative impact on the high-skilled real wage.

Finally, at low trade costs ($\phi = 0.60$, or $\tau = 1.1856$), the symmetric equilibrium becomes unstable and a core-periphery structure arises.

Fig. 1

Low-skilled workers real wage differential: effects of low and high-skilled migrants at different level of trade costs

(a) $S = 1/2$

Regional share of low-skilled workers (U)

(b) $U = 1/2$

Regional share of high-skilled workers (S)

Fig. 2

High-skilled workers real wage differential: effects of low and high-skilled migrants at different level of trade costs

(a) $S = 1/2$

Regional share of low-skilled workers (U)

(b) $U = 1/2$

Regional share of high-skilled workers (S)
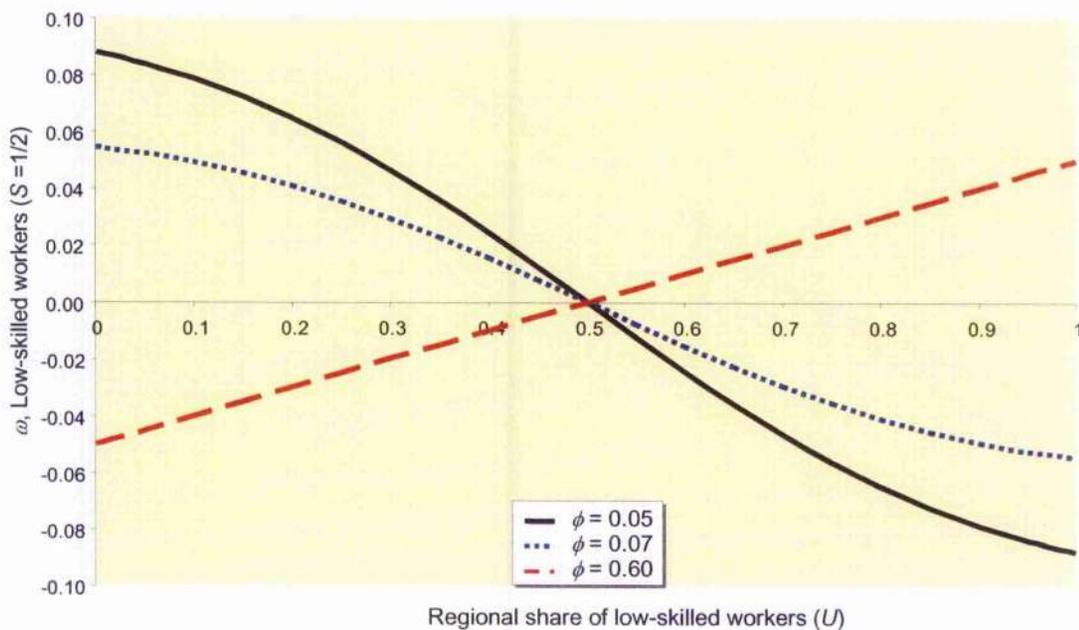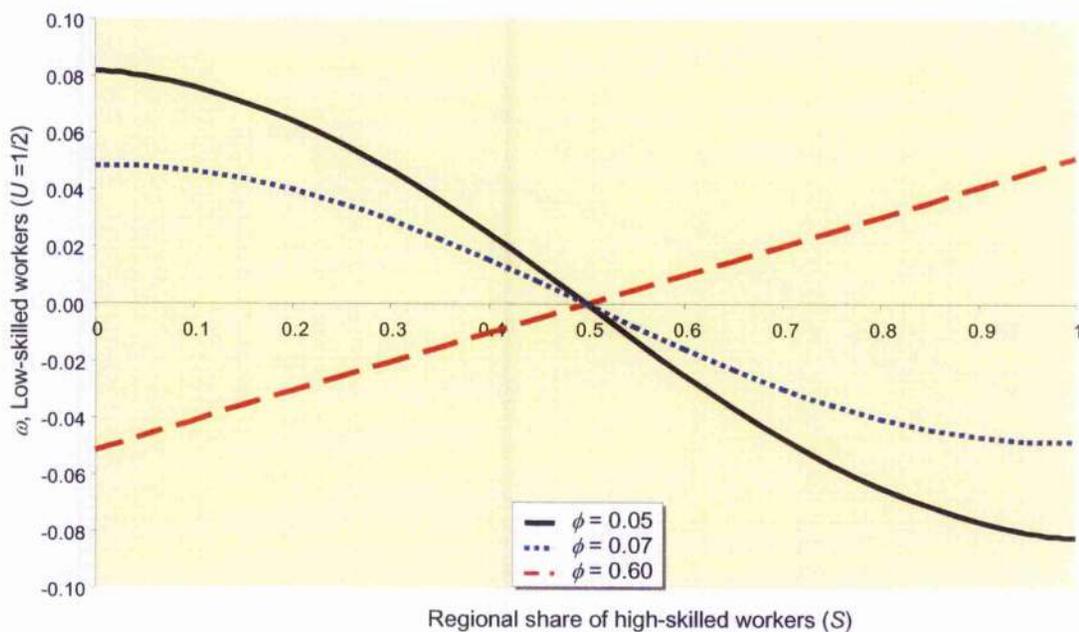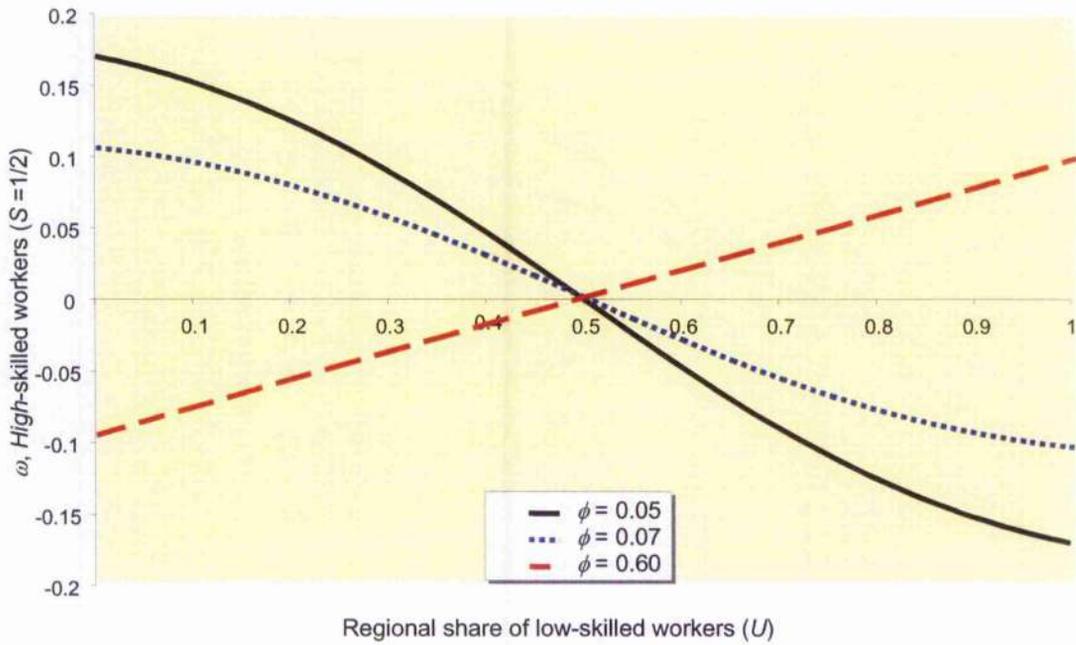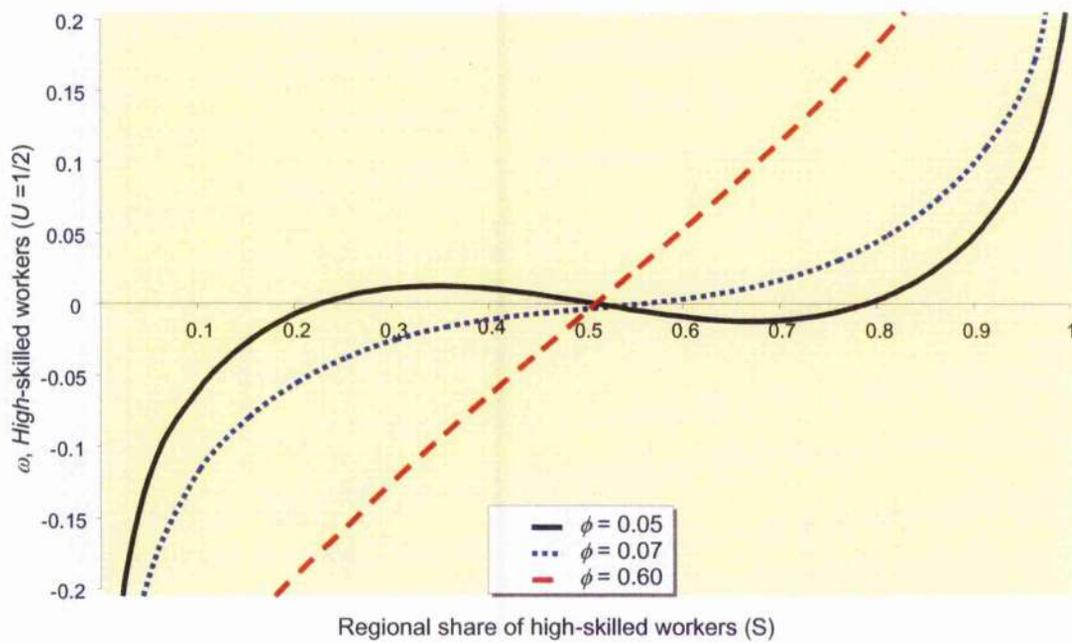
Figure 3 shows how the types of equilibria are related to trade costs. The shares of manufacturing labour force in region I, U and S, are measured on the vertical axis. Solid and dotted lines represent respectively stable and unstable equilibria. At high trade costs, there is a unique stable equilibrium in which skilled and unskilled workers are equally divided between the two regions.

When trade barriers fall below a critical level *(ϕ^s sustain)* a core-periphery equilibrium for the high-skilled becomes possible. However, the symmetric equilibrium is still stable for values above another critical value *(ϕ^s break).*[23] As the regional economies become more integrated high-skilled workers will migrate first in response to real wage differentials.

Is there a range of transport costs for which only the high-skilled migrate in the core region? We compute a critical level of trade costs for which $U \neq 1$ and $S = 1$ (or zero) is a stable equilibrium, *ϕ-self selection.* For trade costs below this level a core-periphery equilibrium as in Krugman (1991) is the only stable outcome. The stability of the symmetric equilibrium for the low-skilled is affected by high-skilled workers' migration. When dispersion forces are still strong (i.e. for high and intermediate level of trade costs) the concentration of high-skilled workers in the core region will push down the low-skilled real wage in the core and therefore cause an outflow of low-skilled from this region until wages are equalized.[24] For sufficiently high trade costs, this equilibrium characterized by highly-skilled concentration in the core and a more than proportional share of low-skilled in the periphery may still be stable. But as regional integration proceeds, the stability may be reversed. A two-stage migration process may arise. First, the core region (which is still

---

[23] For values of trade costs within the range (ϕ^s break < ϕ^s < ϕ^s sustain), both concentration of the high skilled in one of the two regions and symmetry are possible equilibria. The model does not predict which equilibria will arise; policy and history have clearly a potential role in determining the outcome.

[24] This means that once the highly-skilled find it profitable to migrate into the core region a symmetric equilibrium is not sustainable for the low-skilled. This out-migration of low-skilled from the core, often observed in booming urban areas, will reinforce the regional skill differential.

relatively small) is attractive only to skilled and talented individuals. Concentration of skilled workers increases productivity and firms are willing to pay higher nominal wages. A brain drain from the periphery will take place which is both cause and consequence of regional wage disparities. Average wages between the core and the periphery will differ for two main reasons: (i) higher average skills due to sorting of highly skilled in the core; (ii) higher productivity due to localised human capital externalities. Second, as a threshold of trade costs is reached, ϕ-self selection, high-skilled workers' concentration in the core region also induces low-skilled workers to migrate leaving the periphery de-industrialised (the home-market and price-index effects become stronger than the competition effect).

Fig. 3
Trade costs and stability of the spatial equilibria

TRADE LIBERALIZATION
τ decreases
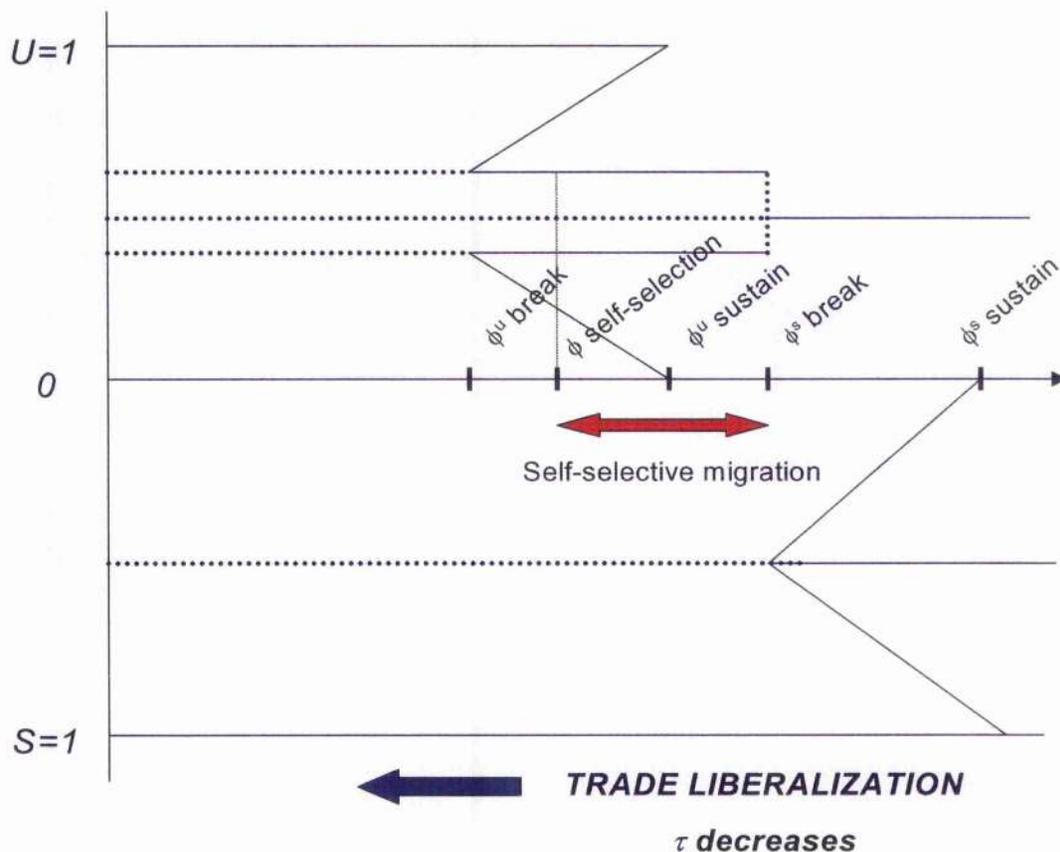
From figure 3 the following relations are evident:

$$\phi^I break < \phi\ self\text{-}selection < \phi^I\ sustain < \phi^S\ break < \phi^S\ sustain$$

there exist, therefore, a range of trade costs for which a self-selecting equilibrium is possible ($\phi\ self\text{-}selection < \phi < \phi^S\ sustain$) and another smaller range for which it is the only stable equilibrium ($\phi^I\ sustain < \phi < \phi^S\ break$).

The critical values considered depend on the parameters of the model. All the threshold values are increasing in μ. If the share of manufactures in the economy is large, the agglomerative forces will be stronger for two reasons. First, the manufactured products will have a bigger share in the bundle of consumption of a representative individual so the *price-index effect* becomes more important. Second, the share of manufactures in world income will be larger and therefore strengthens the *home market effect*.

Critical values are decreasing in σ, the elasticity of substitution in demand. The range of trade costs in which the core-periphery equilibrium occurs is greater the smaller is the elasticity of substitution between varieties. Lower σ implies that consumers view different varieties as being more distinctive, increasing their love of variety. The result is an equilibrium with more varieties and a lower output of each. By decreasing σ, the magnitude of scale economies, one of the agglomerative forces in the model is reduced. Finally an increase in the strength of the positive externalities between high-skilled worker, λ, makes agglomeration possible at higher level of trade costs.

5. *A note on the Beneficial Brain Drain Hypothesis*

A considerable and persistent out-migration of highly skilled individuals has long been thought to be one of the major cause of poor countries/regions remaining poor. Contrary to this widespread belief, a small literature that emphasises a possible positive effect of the brain drain has emerged. The key idea of this strand of research is that, compared to a closed economy, an economy open to the prospect of migration differs not only in the opportunity faced by workers but also in the structure of incentives they face. Higher prospective returns from human capital in the potential destination positively affect the human capital formation decision at home. A positive probability of getting a job abroad might alters individuals' incentives to acquire human capital such that the source economy will end up having an average level of human capital higher than when the probability of migration is zero. This finding is conditional on the fact that not all the individuals succeed in obtaining employment abroad. Stark (2004), for instance, argues that a controlled and restrictive policy can be welfare increasing and might push the economy toward the socially optimal investment in human capital.[25]

The first paper that formalises the effects of migration on levels of education in the source country is Mountford (1997). The author analyses the interaction between income distribution, human capital accumulation and migration. He shows that when educational decisions are endogenous and migration is not a certainty, a brain drain can benefit the sending economy by increasing the proportion of educated people in the economy. Under

---

[25] This result is not possible in a closed economy where individuals under-invest in human capital formation since they do not take into account the social returns from it.

certain conditions[26], the negative effects due to the migration leakage can be outweighed by the increase in human capital formation.

Along the same line Stark et al. (1997) compare the human capital formation and depletion in an economy open to out-migration to that of a closed economy. As in Mountford (1997) the possibility to migrate gives an incentive to acquire human capital at home. The brain gain is ensured not by a strictly positive (and random) probability of successful migration as in the Mountford's model but by imperfect information of employers in the destination country on the skills of the migrants and the presence of return migration (imperfect screening). Heterogeneous individuals (which they label skilled high-skill workers and skilled low-skilled workers) faced with the opportunity to migrate invest in human capital at home and leave the country. Employers in the destination country will initially pay all workers the same wage based on the average productivity. In the subsequent period, employers learn about the true ability of migrants and are able to wage discriminate. The wage adjustment induces the relatively low-skill workers to return in the source country. Under certain conditions the post-return average level of human capital is higher than that of a closed economy.[27] In Stark et al. (1998), a strictly positive probability of migration raises the level of human capital by providing the incentives to invest in education and might therefore result in a brain gain. They use a model in which agents live for two periods. In the first period they split their time between work and education. Work is remunerated by competitive wages per efficiency unit of labour and the cost of acquiring education is equal to the foregone earnings. Investments in human capital increase the efficiency level of labour in the second period. Higher prospective returns from human capital in the foreign country

---

[26] These conditions are a low probability of migration, high returns per efficiency unit of labour in the world economy relative to the home economy and a low proportion of educated people in the pre-migration equilibrium.

[27] In this model the brain gain is more likely to occur: (1) the larger the share of low-skilled workers in the occupation; (2) the steeper the wage schedule with respect to skill in the destination country; and (3) the flatter the wage schedule with respect to skill in the source country.

affect the education decisions made at home. The authors argue that while some workers will leave the country, other workers who have invested in human capital do not succeed in migration and stay home with more human capital than they would have formed in the absence of the possibility of migration.

How likely is this theoretical possibility of a beneficial brain drain for the source country/region? We have reason to believe that, despite its appeal, the likelihood of a beneficial brain drain is rather low. Firstly, the hypothesis of a beneficial brain drain is crucially dependent on a positive and random probability of not being able to migrate. In the context of internal migration in developed countries this optimistic result is more remote since immigration law plays no role. Secondly, the probability of migrating largely depends upon immigration policy in the destination country. In general, immigration policies are designed to give preference to a particular target of migrants. These policies vary a lot across countries and over time. The choice of immigration policy affects the characteristics of the migrants the country will receive, therefore, individuals in the origin country having these targeted characteristics are more likely to migrate.[28]

A simple example will show that a beneficial brain drain is only a special case of a more general framework which deserves further investigation. Many developed countries are experiencing a shortage of low-skilled or skilled manual workers, and their immigration policies are specifically tailored to meet these needs. Individuals will look at the different probability of successful migration across skill profiles and make their human capital investment decision accordingly. Let us suppose we have two countries in the world: Norway and Vietnam. Norway experiences a labour shortage in its fish industry and therefore the government tailors migration policy accordingly. If you are a Vietnamese fisherman you have

---

[28] The only immigration policy which satisfies the assumption of these models, to our knowledge, is the so called *Diversity Visa Lottery* in the US. Only a small fraction of all visas issued in the US is granted through this lottery mechanisms (55 thousands per year) and by contacts of the author with the US Department of States

a probability of migration $p^f = 0.6$, while for any other skill profile the Norwegian government is more restrictive $p^v = 0.2$. In Vietnam, you might either become an engineer with a training cost $k = 0.5$ and earn $W_{ve} = 2$ or remain an unskilled fisherman and earn $W_{vf} = 1$. In Norway fishermen are paid $W_{nf} = 3$ while engineers $W_{ne} = 4$. Given this setting, it is possible to build a simple payoff matrix and analyse human capital formation choice in Vietnam with and without migration.

Table 1 - Payoff matrix: skill formation and location

Skill	Engineer	Fisherman
Location		
Vietnam	$W_{ve} - k = 2 - 0.5 = 1.5$	$W_{vf} = 1$
Norway	$p^v W_{ne} - (1 - p^v)W_{ve} - k =$ $0.2*4 + 0.8*2 - 0.5 = 1.9$	$p^f W_{nf} - (1 - p^f)W_{vf} =$ $0.6*3 + 0.4*1 = 2.2$

With no possibility of migration, to become an engineer is a dominant strategy for a Vietnamese. When there is a positive possibility of migration, the selective Norwegian migration policy might affect the structure of incentives in a perverse way, reducing human capital formation in Vietnam. Clearly this example is too simplistic, nevertheless it gives some insights and stresses the importance of immigration policy in the destination country, a policy variable out of the control of LDCs.

---

it seems evident that a considerable share of the applicants is constitute by illegal migrants or temporary migrants already settled in the US.

6. Concluding remarks

Modern economies have assisted an epochal transition from a world where the basic source of value, productivity and economic growth has been physical labour and manual skill, to a world of knowledge-intensive capitalism where a fundamental role is played by intelligence and intellectual labour (human capital). Scientists, engineers and high skilled workers on the factory floor and in advanced services companies are the sources of innovations and ideas. The *wealth of regions* and nations in the 21st century economies is largely based on this human infrastructure (Florida, 1995). Consequently, a region with a more skilled labour force is likely to grow faster than a region with a less skilled labour force.

Technical progress and reduced barriers to the movement of goods and services have a fundamental impact on the incentive for factor of production to spatially relocate. As emphasised by brain drain literature, migration of high skilled workers may potentially have consequences on the regional economic performance, since it affects the way region maintain and sustain human resources.

This paper uses a version of the core-periphery model of trade and location by Krugman (1991) to analyse the migration behaviour of workers with different skills in a process of regional economic integration. In our model, the interaction between high-skilled workers in a region increases the productivity of each worker by means of a knowledge diffusion process. Regions with a higher average level of human capital are therefore more attractive for high-skilled workers. The skill premium is endogenous and increasing in the regional quota of high-skilled workers.

We find that there exists therefore a range of trade costs for which a self-selecting equilibrium is possible and another smaller range for which is the only stable equilibrium. One of the results of the model is the existence of a persistent wage and productivity

differential between the core and peripheral regions. Recent empirical works have supported this prediction using micro data.

Although in our model migration of highly skilled individual has an impact on low-skilled footloose workers through agglomeration and dispersion forces, in order to keep the analysis simple, we abstract from direct spillovers or complementarities between high- and low-skilled workers in the productive process. A possible extension of the current framework would be that of explicitly taking into account complementarities between high- and low-skilled workers. In addition, the inclusion of congestion costs associated with agglomeration as in Helpman (1997), or of location preferences which translates into an imperfectly mobile labour force as in Ludema and Wooton (1997), would at the same time add more realism to the migratory behaviour and prevent the model from having an unrealistic catastrophic agglomeration at low transport costs.

References

Becker S., Ichino A. and G. Peri (2003), How Large is the Brain Drain from Italy?, manuscript.

Blanchard O., Katz, O.J.L. (1992), Regional evolutions, *Brooking Papers on Economic Activity*, Vol. 1, pp. 1-75.

Borjas G.J., Bronars S.G. and S.J. Trejo (1992), Self-Selection and Internal Migration in the United States, *Journal of Urban Economics*, Vol. 32, pp. 159-185.

Burda M., Wyplosz C. (1992), Human Capital, Investment and Migration in an Integrated Europe, *European Economic Review*, Vol. 36, pp. 677-684.

Chiswick B. R. (2000), Are Immigrants Favourably Self-Selected? An Economic Analysis, IZA Discussion Paper N. 131.

Ciccone A., Peri G. (2000), Human Capital and Externalities in Cities, CEPR Discussion Paper No. 2599, London.

Ciccone A., Hall R. (1996), Productivity and the Density of Economic Activity, *American Economic Review*, Vol. 86, pp. 54-70.

Combes P. P., Duranton G. and Gobillon L. (2003), Spatial Wage Disparities: Sorting Matters!, manuscript.

Commander S., Kangasniemi M. and Winters L.A. (2004), The Brain Drain: Curse or Boon? A Survey of the Literature, in Baldwin R. E. and L. A. Winters (eds), *Challenges to Globalization: Analysing the Economics*, The University of Chicago Press, Chicago.

De la Fuente A., Donenech R. (2000), Human Capital in Growth Regressions: How Much Difference does Data Quality Make?, *OECD Working Paper* No. 262.

Decressin J., Fatas A. (1995), Regional Labour Market dynamics in Europe, *European Economic Review*, Vol. 39, pp. 1627-1655.

Eaton J., Eckstein Z. (1997), Cities and Growth: theory and evidence from France and Japan, *Regional Science and Urban Economics*, Vol. 27.

Fujita M., Krugman P. and Venables A. J. (1999) *The Spatial Economy: Cities, Regions and International Trade*. The MIT Press, Cambridge.

Florida R. (1995), Toward the Learning Region, *Futures*, Vol. 27 (5), pp. 527-536.

Giannetti M. (2003), On the Mechanics of Migration Decisions: skill complementarities and endogenous price differentials, *Journal of Development Economics*, Vol. 71, pp. 329-349.

Glaeser E.L. (1998), Are cities dying?, *Journal of Economic Perspectives*, Vol. 12 (2), pp.139-160.

Glaeser E.L., H. Kallal, J. Sheinkman, and A. Shleifer (1992), Growth in Cities, *Journal of Political Economy*, Vol. 100, pp. 1126-1152.

Hanson G. (2001), Scale Economies and the Geographic Concentration of Industry, *Journal of Economic Geography*, Vol.1 (3).

Harris, C.D. (1954), The Market as a Factor in the Localization of Industry in the United States, *Annals of the Association of American Geographers*, Vol. 44, pp. 315-348.

Helpman E.(1997), The Size of Regions, in *Topics in Public Economics. Theoretical and Applied Analysis*, D. Pines, E. Sadka and I. Zilcha (eds.), Cambridge University Press, Cambridge.

Kremer M. (1993), The O-Ring theory of Economic Development, *Quarterly Journal of Economics*, Vol. 108 (3), pp. 551-575.

Krugman P. (1991) Increasing returns and economic geography, *Journal of Political Economy*, Vol. 99, pp. 484-499.

Lucas R. E. (1988), On the mechanics of economic development, *Journal of Monetary Economics*, Vol. 22, pp. 3-42.

Ludema R. and Wooton I. (1999), Regional Integration, Trade, and Migration: Are Demand Linkages Relevant in Europe?, in *Migration the Controversies and the Evidence*, Faini R., de Melo J. and K. Zimmermann (eds.), CEPR London.

Moretti E. (2004), Humna Capital Externalities in Cities, in V. Henderson and J-F. Thisse (eds.), *Handbook of Regional and Urban Economics*, volume 4, North-Holland (Amsterdam), forthcoming

Mori T., Turrini A. (2004), Skills, Agglomeration, and Segmentation, *European Economic Review*, forthcoming.

Mountford A. (1997), Can a brain drain be good for growth in the source economy?, *Journal of Development Economics*, Vol. 53, pp. 287-303.

Neary, J. P. (2001), Of Hype and Hyperbolas: Introducing the New Economic Geography, *Journal of Economic Literature*, Vol. 39 (2), pp. 536-561.

Ottaviano G. and Puga D. (1998) Agglomeration in the global economy: a survey of the new economic geography, *The World Economy* 21, pp. 707-731.

Overman H., Puga D. (1999), Unemployment Clusters across European Region and Countries, CEPR Discussion Paper No. 2225, London.

Policy Options, The Brain Drain, Issue September 1999, Institute for Research on Public Policy, Montreal (Quebec), www.irpp.org/

Puga D. (1999), The Rise and fall of Regional Inequailities, *European Economic Review*, Vol. 43, pp. 303-334.

Quah D. (1996), Regional Convergence clusters across Europe, *European Economic Review*, Vol. 40.

Rauch, J.E. (1993), Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities, *Journal of Urban Economics*, Vol. 34, pp. 380-400.

Rauch, J. E. (1999), Networks vs. Markets in International Trade, *Journal of International Economics*, Vol. 48 (1), pp. 7-35

Roback J. (1982), Wages, rents, and the quality of life, *Journal of Political Economy*, Vol. 90, pp. 1257-1278.

Rosenthal S., Strange W. (2004), Evidence on the nature and sources of agglomeration economies, in V. Henderson and J-F. Thisse (eds.), *Handbook of Regional and Urban Economics*, volume 4, North-Holland (Amsterdam), forthcoming

Saxenian A., Y. Motoyama and X. Quan (2002) Local and Global Networks of Immigrant Professionals in Silicon Valley, Public Policy Institute of California, San Francisco (CA).

Stark O., Helmenstein C., Prskawetz A. (1997), A Brain Drain with a Brain Gain, *Economic Letters*, Vol. 55, pp. 227-234.

Stark O., Helmenstein C., Prskawetz A. (1998), Human capital depletion, human capital formation, and migration: a blessing or a "curse"?, *Economic Letters*, Vol. 60, pp. 363-367.

Stark O. (2004), Rethinking the Brain Drain, *World Development*, Vol. 32 (1), pp. 15-22.

Sveikauskas, L. (1975), The Productivity of Cities, *Quarterly Journal of Economics*, Vol. 89, pp.393-413.

Chapter 3

Human Capital Accumulation and Migration of the Highly Skilled in a Peripheral EU Region: the Case of Basilicata[°]

Nicola D. Coniglio

1. Introduction

The importance of human capital as a key source of value added, innovation and economic growth is widely acknowledged by both economists and policymakers. In his seminal paper "Reflections on Investment in Man", Schultz (1962) underlines the importance of improvements in the quality of human capital as one of the major factors explaining economic growth. The abilities of individuals, "the economic capabilities of man", are considered as *"produced means of production* and except for some pure rent (in earnings) for differences in inherited abilities, most of the differences in earnings are a consequence of differences in the amounts that have been invested in people" (Schultz, 1962, p.1). The abilities of individuals are, therefore, the product of individual investments in education and job-training activities.

The relevance of human capital in modern economies, and the divergence between its private and social return, justifies public policy interventions aimed at subsidising human capital accumulation.[1] The regional dimension is increasingly becoming central in this process of knowledge creation: local governments can directly affect individuals' decision making by subsidising human capital formation. However, as Justman and Thisse (1997) have pointed out, the public benefits of human capital investment "do not always accrue where it is funded". The scope for policy might vanish if locally formed human capital is lost through migration, as emphasised in the *brain drain* literature. Positive self-selection is a stylised fact in the migration literature, i.e. more talented and skilled individuals have an higher propensity to migrate than the rest of the population. This can result in migration widening the regional wage and income gap and lowering the standard of living in the

---

[1] Educational expenditure is considered a key component of national investment with a substantial pay-off in terms of output growth. The current state of thinking about this issue is well summarized by Temple (2001), who, after surveying the relevant micro- and macroeconomic evidence, concludes that "the weight of the evidence points to significant productivity effects of educational investment".

peripheral location, as the new economic geography literature emphasises (Krugman, 1991; Coniglio, 2003). Regional integration of labour markets may increase workers' incentive to spatially relocate, in particular from peripheral to central regions. This is explained by the action of agglomeration forces: human capital migrates from where it is scarce to where it is abundant, rather than vice versa (Lucas, 1988).

In the present study, we emphasise the importance of these considerations through the analysis of a case study. We investigate the location decision of a sample of highly educated and skilled individuals residing in Basilicata, a small peripheral Italian Mezzogiorno region, who have benefited from a locally funded human capital investment policy. The regional policymakers, in recognition of the importance of human capital as a key ingredient for regional growth, have given generous subsidies since the beginning of the 1990s to young graduates who want to attend postgraduate courses either inside or outside the region.

The sample of individuals under analysis is far from being representative of the regional population, but the biased nature of the sample (highly skilled and educated individuals) makes the analysis a natural experiment for assessing the ability of the regional system in a peripheral region not only to *generate* human capital but also to *maintain* it.

The focus of the paper is on the micro-level migratory behaviour. In particular, we want to shed light on the following questions:

1) Who are the migrants? We want to investigate whether there is a significant difference in the personal characteristics of migrants and non-migrants.

2) What are the main factors influencing migration?

3) Where do migrants go? Is human capital attracted toward core regions? What are the regional characteristics that make a region attractive for highly educated and talented individuals?

55

The data set used in this analysis has been generated through a postal survey designed and conducted by the author. The survey questionnaire was sent to all the individuals who benefited from the local policy measure during the period 1991 to 2001. For each individual, we collected data on background, experience and outcome of higher education, opinion on the quality of the course attended, job-search strategy after the course, and space-time career details from first to current employment status. The high quality of the data and the extremely high rate of response (around 70%) makes the data set at our disposal a unique tool for studying the micro-level migration decision of highly skilled individuals in a peripheral region.[2]

The empirical analysis aimed at addressing the above-mentioned questions consists of two main parts. In the first part, we investigate the personal characteristics explaining the propensity whether or not to migrate. This analysis exploits maximum likelihood estimation with a logit model. We are able to study different propensities to migrate within a particularly interesting stratum of the population, young and highly educated individuals. Prior insights from migration literature tell us that this sub-group of the population is highly mobile, but to our knowledge there are no in-depth studies focused on their micro-level migration decision. We find that, even in this homogeneous sample, individuals show a different migration propensity on the basis of their personal characteristics. The probability to migrate decreases with age and there is a positive self-selection in migration flows: even among a group of highly educated individuals, the probability to move is higher for the most talented. Individuals with different educational backgrounds have a different propensity to migrate: an individual with a bachelor's degree (or master's degree) in business studies or engineering is less likely to stay home than other graduates. An important finding of our analysis is that individuals who studied outside the area of origin (that is individuals with a previous

---

[2] Detailed descriptions of the survey methodology, the questionnaire design and of sample selection issues are

migration experience) are more likely to migrate. It is worth noticing for its policy implications, that, attendance at a post-graduate course, or getting an internship, at home significantly increases the probability of retaining the highly skilled worker. We find also that the effect of the personal unemployment status on the propensity to migrate is negative, that is, unemployed individuals are mainly located in the home region. This gives support to the "contracted" view of migration, i.e. individuals migrate with "a job in their hands". Moreover, people born in urban areas are less likely to migrate compared to those born in small towns.

While in the first part of the analysis we investigate the personal characteristics which make an individual more or less likely to migrate, in the second part we analyse an important complementary issue, namely the characteristics which make a region attractive to highly skilled individuals. In the second part of the paper we jointly analyse the decision to migrate and the choice of the location of migration, i.e. we investigate the spatial structure of the human capital generated through the policy intervention. We use a conditional logit model in order to investigate where highly educated and skilled individuals do go. The use of a conditional logit model for the study of individual migration decision has considerable advantages, since it enables the researcher to take into account the opportunities that individuals face in a set of potential alternatives. In the present study, each Italian region is considered as a potential destination (including Basilicata). Individuals choose the location which offers the higher level of utility on the basis of regional and personal characteristics.

Knowledge of the factors influencing highly skilled individuals' location decisions is fundamental to national and local policymakers who wish to capitalise on this important source of economic development. We study the decision made by both movers and non-movers. Thanks to the availability of very detailed information on individuals' characteristics

---

reported in chapter 4 of this thesis.

and past experience, we are able to assess the relative importance of some variables like distance and the individual's previous migration experience in a way that is generally not possible due to data limitations.

We find a strong evidence for the importance of spatial variables, such as physical distance, past migration flows from the origin region and direct knowledge of potential destinations due to previous migration experiences, in explaining the pattern of subsequent geographical mobility. Lack or poor quality of information on local economic opportunities represents a market failure that can be, at least partly, corrected by adequate policy measures. Action aimed at this purpose might also be considered as strategic in trying to maintain and attract human capital. Moreover, non-economic aspects are important in influencing high-educated individuals' location choice. Local policy aimed at improving quality of life should be regarded as an important policy instrument along with more traditional interventions.

Our analysis explores the factors affecting the location decision of highly educated individuals in an economy presenting all the features of a core-periphery structure.[3] Our results have strong policy implications for the local economy under analysis. However, in our opinion, they also have a general validity for regions facing similar circumstances, such as many peripheral regions across Europe.

The structure of the paper is as follows. In section 2, we describe in more detail the regional policy implemented, the conduct of the survey and the sample under analysis. In section 3, the logit model is presented and the results are discussed. The conditional logit analysis is presented in section 4. Finally, section 5 contains our concluding remarks on the paper.

---

[3] Of course, as with all case studies, it is necessary to exercise caution when generalising the findings to other situations. Nevertheless, it is reasonable to argue that the behaviour of the highly educated and skilled individuals in our sample is not substantively different from that of similar individuals in peripheral regions of many developed countries.

58

2. Conduct of the survey and description of the data

The data used for the empirical analysis are derived from a postal survey of individuals who have received financial support from Regione Basilicata for attending a postgraduate course in a Higher Educational Institution.[4] The survey was designed and conducted by the authors with the collaboration of Regione Basilicata. The survey questionnaire aims at directly asking people about their decision whether or not to move and the main factors influencing their decision.[5] For each individual, we have collected data on background, experience and outcome of higher education; opinion on the quality of the course attended; job-search strategy after the course; and space-time career details from first to current employment status. The survey was conducted during May-July 2002. The questionnaire was sent to all the individuals who benefited from the local policy measure between 1991 (the first year in which this measure was implemented) and 2001, which amounts to slightly more than one thousand individuals.[6] A total of 740 individuals returned the questionnaire, approximately 70% of the total. After deleting observations relating to incomplete questionnaires, the final sample contains 700 observations.

The individuals in our sample have a bachelor's degree and were resident in Basilicata at the time when they applied to receive funding.[7] The number of migrants at the time of the survey was 411 (58% of the total). It is worth noting that Basilicata, like the rest of the Italian Mezzogiorno, has been a traditional source of migrants for both other Italian regions and foreign destinations. The propensity to out-migrate toward other Italian regions as measured

---

[4] Postgraduate courses were held in Basilicata or outside (also abroad). In what follows we use interchangeably the words postgraduate course and master's degree for identifying the human capital investment experience financed by the regional authority.

[5] In what follows we investigate individuals' migration decisions during the period 1991-2002.

[6] In order to increase participation, individuals received a follow-up request letter and subsequently were contacted by phone.

[7] There is a small group of people (35 individuals) whose highest qualification at the time of application was a high school diploma.

by the ratio between the (gross) number of people changing their registered place of residence and the total regional population (multiplied by 1.000), is considerably higher than the Italian average. In 1997, this index of migration was 8.4 for the overall Basilicata population against 5.3 for the Italian regional average and 7.2 for the Italian Mezzogiorno (ISTAT, Italy's National Institute of Statistics).

Naturally, we should expect that young individuals from Basilicata have a higher propensity to migrate. ISTAT's data reveals that in 1997, of the cohort of 5,137 individuals aged 25-29 in Basilicata, 990 changed their residence to other Italian regions. For this cohort the migration index, as defined above, was 22.6 for male and 18.7 for female (the Italian average was 11.1 for males and 9.9 for females).[8]

3. Who are the migrants?

3.1. Empirical Methodology

A first step of the empirical analysis is related to the discrete binary choice of an individual whether to remain in the region of origin or move to another location. The econometric analysis of this study is based on the micro-level decision-making of a potential migrant. Formally the level of utility enjoyed by an individual i at the present location (h) is:

$$U_h^* = \beta_h ' x_h + \varepsilon_h \tag{1}$$

---

[8] Note that in our sample the average age is 27.5 with a standard deviation of around 2.5 years.

where x_h is a vector of independent variables and β' the associated coefficients measuring the relative importance of these variables. This vector includes personal characteristics such as age, sex, education and variables reflecting regional characteristics of the origin area.

If the individual migrates to another location (f) the level of utility she will enjoy is given by:

$$U_f^* = \beta_f ' x_f + \varepsilon_f \tag{2}$$

The opportunities and costs and, therefore, the associated level of utility which the individual will face are a function of personal characteristics and regional characteristics of the destination area. In the balance between these two levels of utility the effects of intervening obstacles should be taken into account. Costs of migration are a function of individual characteristics and existing barriers:[9]

$$CM_{hf}^* = \alpha ' x_c + \varepsilon_c \tag{3}$$

An individual remains in the current location if the benefits from migration $U_h^* - U_f^*$ are smaller than the associated cost CM_{hf}^*:

$$
\begin{aligned}
H^* &= U_h^* - U_f^* + CM_{hf}^* = \beta_h ' x_h - \beta_f ' x_f + \alpha ' x_c + \left(\varepsilon_h - \varepsilon_f + \varepsilon_c \right) \\
&= \beta ' x + \varepsilon
\end{aligned}
\tag{4}
$$

---

[9] Barriers to labour mobility and migration can be of different nature. Factors such as for example geographical distance, uncertainty linked to a lack of information, a different environment, legal barriers, are all important and might be perceived in a different way by heterogeneous individuals.

where H^* is an unobservable variable representing the net benefit from remaining in the current location. The observed binary individual choice reveals only if the net benefit is positive or negative, but not the underlying unobservable utilities. Our observation will be:

$H = 1$ if $H^* > 0$

$H = 0$ if $H^* \leq 0$

On the basis of the outlined random choice approach, it is possible to examine the probability that an individual with a given set of attributes will make the choice of staying in the current location ($H = 1$) or migrating ($H = 0$):

$$\text{Prob}\left[H = 1 | x\right] = \text{Prob}\left[U_h^* - U_f^* + CM_{hf}^* > 0\right]$$
$$= \text{Prob}\left[\beta' x + \varepsilon > 0 | x\right] \tag{5}$$

We assume that the disturbance ε has a logistic distribution with mean 0 and variance 1. This assumption will lead to a logit model.[10]

3.2. Description of the data and variables influencing migration

The dependent variable **HOME** is a dummy that equals 1 if the individual resides in Basilicata at the time when she received the questionnaire and 0 otherwise. We look at individuals' current locations.[11]

---

[10] Logit and probit models are widely used in econometric applications. The difference between the two models is based on the particular probability distribution function adopted. While the former uses a logistic distribution function the latter is based on a cumulative normal one. The two distributions are similar except for the heavier tails of the logistic distribution. On theoretical grounds it is difficult to justify the choice of one instead of the other, since it makes very little difference in terms of substantive results.

The essence of our approach is to determine the characteristics that differentiate the two groups of individuals and to evaluate the impact of some individual characteristics on the propensity to migrate.

The explanatory variables used in the econometric model are defined in *table 1*. The set of variables include (*i*) individual characteristics such as age, sex, city of origin and employment status (*ii*) detailed information on educational background (marks obtained, subject of study, location of university and master studies, eventual internship period after the master, etc.) (*iii*) information on geographical preferences in the job-market search after completion of the postgraduate course.

The literature on migration has shown that among the factors which contribute to the decision to migrate a relevant role is played by personal characteristics (sex, age, etc.) and a number of life-cycle considerations (marriage, completion of schooling, entry into the labour force, etc.).

Several studies have shown that the peak migration propensity occurs in the early twenties and declines steadily thereafter (Ledent, 1990; Otomo, 1990; Vergoossen, 1990; Long, 1992; Plane, 1993). We investigate the relationship between age and migration propensity by using the variables *age* and *age squared*.[12] Age of each individual is measured at the moment when the bachelor's degree was completed. In our specification this variable should be interpreted as a qualitative measure of personal abilities rather than being aimed at capturing life cycle effects. The expected sign for *age* is positive since younger individuals are those who completed the university studies in a shorter time span, i.e. the most talented individuals.

---

[11] In section 4, we estimate a conditional logit model to examine the location decisions of these individuals at completion of their master studies. In that section we discuss the motivations which lead us to investigate individuals' location decisions in two different points in time.

[12] Age squared is included to test whether there is a non-linear effect of age on the dependent variable.

The relationship between education and migration propensity has been widely studied in empirical works. High educational attainments are generally employed as a proxy for individual abilities. Since more talented and able individuals have both lower costs and higher returns from migration, education is positively associated with migration probability (Ghatak *et al.*, 1996; Antolin and Bover, 1997; Ritsila and Ovaskainen, 2001). This implies that migrants tend to be favourably "self-selected" for labour market success (Borjas, 1987; Borjas *et al.*, 1992a,b). On the basis of these characteristics, individuals in our sample should be considered as highly mobile. We want to test if, even within this sample, variation in these personal characteristics matters in explaining whether individuals are likely to stay or move. In our study we employ a more direct measure of educational attainment, i.e. the final mark of the bachelor's degree (*mark*).

The expected effect of gender on the propensity to migrate is ambiguous. Traditionally male individuals have been found to be more likely to migrate, but, as stated in Ghatak *et al.* (1996), several studies have shown that this is no longer typical.

The effect of the size of the place of origin on the propensity to migrate is not well documented in the literature. Is there any significant difference in the propensity to migrate between individuals born in larger urban areas compared to those born in rural areas or small towns? The variable *city* allows us to test for significant differences between people born in a city and those born in smaller towns.[13] We expected that this variable would have a positive sign, since the cost of migration for individuals born in a city is higher than for residents in smaller centres, which have stronger push factors for young and highly educated individuals.

Beyond these factors, we want to check if individuals with different educational backgrounds might have a different propensity to migrate (*business*, *engineering* and *law*).

---

[13] Individuals born in a city are those born in Matera or Potenza, which are the two main urban agglomerations in Basilicata.

For example, an engineer and a lawyer might face a considerably different spatial distribution of job-market opportunities.

Individuals' past migration experience during the period of investment in human capital is essential for understanding subsequent patterns of geographical location. If an individual has lived in a location different from home during her university or master studies, she is more likely to migrate for two main reasons. First, these experiences enable the building of personal connections (with friends, potential employers, etc.), which reduce the information costs of finding employment opportunities. Moreover, knowledge of the local environment reduces the costs of adapting to new environments. "The [...] repeat migrant [...] have substantial first-hand knowledge about locations where he lived before, and [...] he may be able to recoup some part of any location-specific capital he acquired there" (DaVanzo, 1983). Second, individuals who have already moved once could have lower costs associated with a second move; for example in term of psychic costs, since original family ties and the like have already been broken (Farber, 1978 and Herzog and Schlottmann, 1981).[14] We test for the relevance of these considerations by means of two dummy variables: *university away* and *master away*. According to our hypothesis we expected a negative signs for all thesevariables.

Since an internship was for some individuals an integral part of the master course financed, we include a dummy variable *internship* (equalling 1 if the individual did an internship and 0 otherwise) to investigate whether this individual experience might explain part of the propensity to migrate. In addition, we consider the effects of doing an *internship at home* adopting a dummy variable which is equal to 1 when the master course was held at home and 0 otherwise. In this case the sign of the estimated coefficient should be positive. [15]

---

[14] Molle and van Mourik (1989) found that the psychic costs significantly reduce migrations within Europe.

[15] In this study the Universities of Salerno and Bari are considered as "home" locations. This is justified on the basis of strong economic and social connections between the provinces of Matera and Bari, on one side, and between the provinces of Potenza and Salerno, on the other side. Moreover, Salerno and Bari are close, respectively, to Potenza and Matera and well connected. We also have to consider that University of Basilicata does not have a long tradition, so the Universities of Bari and Salerno attract many students from Matera and Potenza.

Table 1 – Probability of staying in the home region: description of the explanatory variables used in the logit model

Variable	Description	Expected sign of effect	Mean	Std. dev.
Male	1 = male 0 = female	+ / -	0.44	0.49
City	1 = born in city 0 = born in a small town	+	0.57	0.49
Age	Age at the time of attending the postgraduate course	-	27.47	2.75
Age squared	Age, as defined above, squared	+ / -	117.28	63.83
Mark	Final mark of the bachelor's degree (between 66 and 110)	-	102.7	8.17
Business	1 = degree in business studies 0 = other	+ / -	0.36	0.48
Engineering	1 = degree in engineering or architecture 0 = other	+ / -	0.09	.028
Law	1 = degree in law 0 = other	+ / -	0.31	0.46
University away	1 = undergraduate studies attended outisde the home region 0 = other	-	0.48	0.50
Master away	1 = Master course attended outside the home region 0 = other	-	0.75	0.43
Internship	1 = individual obtained and completed an internship 0 = other	-	0.79	0.41
Internship at home	1 = internship at "home" (in Basilicata or in the provinces of Bari and Salerno) 0 = other	+	0.15	0.36
No_pref_HOME	1 = Basilicata is not the preferred location 0 = other	-	0.64	0.48
Unemployed	1 = currently unemployed (at the moment of the survey) 0 = other	+ / -	0.15	0.36
High mobility	1 = individual changed its jobs more than 3 times between the first and the current employment position 0 = other	-	0.11	0.31
Masteryear	The year when master course has been attended	+ / -	7.53	2.21
Highschool	1 = individual's highest degree is highschool diploma 0 = bachelor's degree	-	0.05	0.22
Age*Highschool	Variable capturing the interaction effect between *age* and *highschool*	-	0.35	1.76
Age squared* Highschool	Variable capturing the interaction effect between *age squared* and *highschool*	+	3.22	19.81
Mark*Highschool	Variable capturing the interaction effect between *mark* and *highschool*	-	4.06	17.95

In the questionnaire we asked individuals about the spatial structure of the job-search after completion of the master course. One of the items concerned individuals' spatial-preferences. The variable *no_pref_HOME* captures, *ceteris paribus*, the effect of a lack of personal attachment to the origin region. Obviously, we expected a negative sign for the coefficient of this variable. Implicitly, the baseline model would include individuals with a strong attachment towards Basilicata.

Another personal characteristic we deal with is *unemployment* status.[16] The expected sign of the coefficient associated with this variable the literature is controversial. According to the "contracted" migration view, the sign should be positive, that is unemployed individuals should mainly be located in the home region since individuals migrate with a "job in their hands". Opposed to this view is the idea of "speculative" migration, which sees individual migration as part of the job-search strategy. As pointed out by Herzog *et al.* (1993) in their survey of the empirical literature concerning the complex interrelationships among personal unemployment, migration and the likelihood of re-employment, the availability of microdata is essential to study the relationship between unemployment status and migration.

The variable *high mobility* measures the number of individual's job changes. We expect a negative sign, since individuals who change job frequently should be also more spatially mobile.

Since we are investigating migration decisions along a ten-year time span, it is necessary to control for different propensities to migrate due to changes in economic conditions in the alternative destinations. For this reason we introduce the variable *masteryear* as a control variable.

---

[16] Unemployment might affect mobility at three levels: (1) personal unemployment; (2) regional unemployment differentials; (3) national unemployment rates. Here we analyse only the effect of personal unemployment, since we are studying the personal characteristics explaining the propensity to migrate. In order to consider the effect of the regional unemployment differentials, we included the regional unemployment rates in the conditional logit we estimate in the section 4.

In addition, we include three interaction effects which relate to the small number of individual in our sample whose highest degree obtained is an highschool diploma (*age*highschool, age squared*highschool* and *mark*highschool*). These control variables are necessary since these individuals benefited from the policy at a younger age than the rest of the sample and their final mark is expressed in a different scale which we converted to the scale of the bachelor's degree final mark.

3.3. Discussion of the empirical results

Table 2 reports the results of the econometric analysis. Most of the coefficients are statistically significant and have the expected signs. Since the estimated parameters of a binary regression model do not provide directly useful information for understanding the relationship between the independent variable and the outcome, in the last column the marginal effects are reported.[17]

We find evidence of differences in the migration behaviour between male and female individuals. In our analysis male individuals are more likely to stay home, more precisely the difference in the probability of staying home, evaluated at the baseline group, is 9.2%. This is an interesting result, even because the probability to migrate for the female population in Basilicata, in the cohort aged 25-29, is lower than that for male individuals in the same cohort. This finding points out that for female individuals the propensity to move increases relatively more than males if they are highly educated.

People born in urban areas are 8.5% less likely to migrate. This outcome can be probably explained by the fact that larger cities are characterised by a critical mass of

---

[17] The baseline group is defined as an individual with the following characteristics: a female; born in a urban area; with a bachelor's degree other than economics, law or engineering; attending undergraduate studies at home and master studies in a different locality; internship not at home; with expressed preferences for the home

amenities, therefore, the push factors are smaller as compared to towns.[18] In addition, densely populated areas can offer more job opportunities. Individuals from small centres are likely to move anyway for job reasons; hence, for these sub-groups the substitutability between a home or outside location is higher.

The probability of remaining in the home region increases with age. We find support for a non-linear effect of this variable. Since age is measured at the moment when the bachelor's degree was completed, in addition to the life cycle considerations which link age and migration propensity, this variable might also be interpreted as a qualitative measure of personal abilities. The positive sign of the estimated coefficient confirms that there is a positive self-selection in migration flows. Even among a group of highly educated individuals, the probability of moving is higher for most talented.

Further support for the existence of a positive self-selection of these migrants is given by the high significance of the variable *mark*. Individuals with a high final mark of the bachelor's degree are less likely to stay home. An individual with the lowest mark is 36% more likely to stay at home than an individual with the top mark.

Individuals with different educational background have a different propensity to migrate. Among the variables which describe the subject of the university courses attended, we find that *business* and *engineering* are significant. The negative signs are probably justified by the fact that the "pull" factors for individuals with a degree in engineering and business are higher because some of these individuals are going to be employed in sectors like R&D, banking, etc, which are typically under-represented in peripheral, as opposed to core, regions. Individuals with a first degree in business studies and engineering are 18.4% and 21%, respectively, less likely to stay home than other graduates. This outcome has strong

---

location and currently employed. All continuous variables are specified at the mean value. The probability of staying at home for the baseline individual is 0.477.

[18] For instance, larger cities offer a wider selection of restaurants, cinemas and in general a wider set of consumption possibilities compared to small towns.

policy implications: if policymakers give priority to financing human capital formation in subjects that are more closely connected with the needs of the regional economy this should increase the probability of maintaining the human capital generated.

In our analysis the most important factors are those related to the location in which people attended university, studied for their postgraduate qualification and obtained an internship. Our hypothesis is that individuals who attended their studies outside the area of origin are more likely to migrate. This is explained by the fact that individuals with previous migration experience in a given destination acquire information on the local labour market at lower costs through a series of sources (interpersonal communications with friends, potential employers, local communication media etc.). In addition, non-monetary and psychological costs of migration will be lowered by an increased familiarity with the local environment. The estimation results support our hypothesis. *Master away* and *internship at home* both have a strong marginal effect. Undertaking a postgraduate course outside the region reduces the probability of remaining in the home region by 21.8%. In **figure 1**, we plot the effect on the predicted probability of staying at home of a change in the final mark of the bachelor's degree for individuals who attended the postgraduate course inside or outside Basilicata. The effect of this variable is rather strong.

Figure 1. Predicted probabilities: the effect of mark for individuals who attended a master at home or away
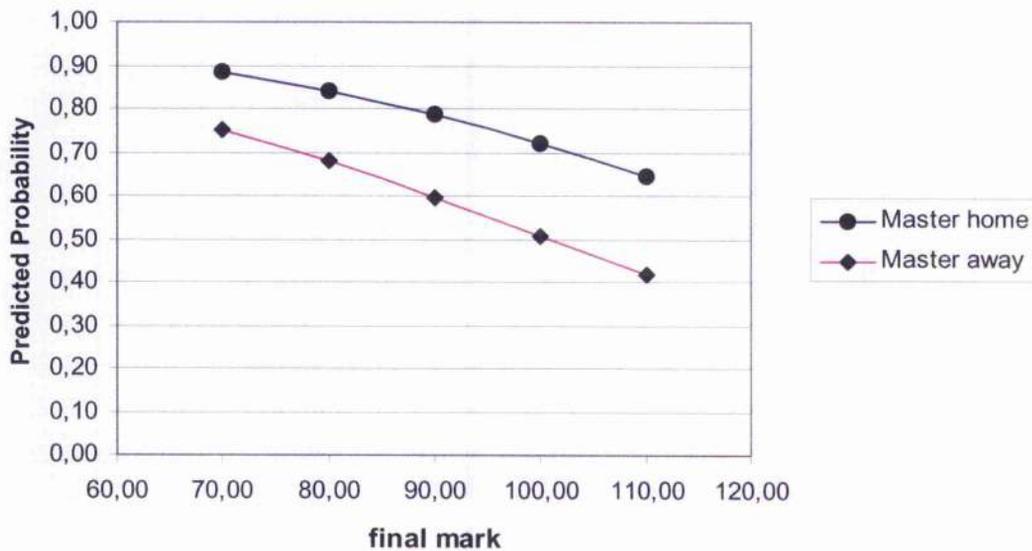
The coefficient of the variable *internship at home* captures the effect on the probability to stay of an internship in the "home" location. Note that the dummy variable *internship* is significant and has a negative estimated coefficient. This outcome confirms our previous hypothesis: undertaking an internship increases the probabilities that an individual will remain in that location. Improving the connections with local economic agents by integrating higher education courses with internships in local enterprises and institutions would improve significantly the ability of retaining human capital. In fact, an internship at home increases the probability of staying by 7.3%.

The variable *no_pref_HOME*, as was expected, has a strong impact (a marginal effect of 22.5%) and is very significant.

The effect of personal unemployment on the propensity to stay home is positive. That is, unemployed individuals are mainly located in the home region. This outcome is in line with the "contracted" migration view, which predicts that individuals migrate with a "job in their hands". It is interesting to compare this finding with the results of other studies using

microdata.[19] In contrast to our results they find that unemployed individuals are more likely to move than the employed.[20] Our results could be explained by the presence of family networks that work like a social security system. Moreover, if we interpret the unemployment status as a proxy for low individual abilities, another explanation could be given by the presence of a process of positive self-selection among the individuals of our sample.

Our analysis confirms the hypothesis that the more dynamic individuals in the labour market are also the more spatially mobile (*high mobility*). These people are 15.6% less likely to remain in Basilicata.[21] The relationship between residential movement probabilities and job changing has been analysed also in Molho (1987), who obtains different results. Molho's study uses a microdata set and is devoted to the analysis of males aged between 16 and 24 in Great Britain. It shows that residential moves are less common amongst those changing occupation. However, the results become similar, when it considers only the individuals changing occupation and improving their "job position".

---

[19] See, for example, DaVanzo, 1978; and Herzog and Schlottmann, 1984, whose geographical focus is on the United States, and Molho, 1987; Hughes and McCormick, 1989; Pissarides and Wadsworth, 1989 and Antolin and Bover, 1997 for Europe.

[20] Some studies have shown that the positive effect of personal unemployment on the propensity to migrate becomes lower the larger the unemployment insurance given (Antolin and Bover, 1997) and the more lasting is the unemployment (McHone and Rungelin, 1993).

[21] Caution should be given to the interpretation of this variable. The causality might actually work in the opposite direction; people residing in Basilicata could be less dynamic as a consequence of local labour market conditions. We need to further investigate this aspect.

Table 2 – Probability of staying in the home region: logit regression results

Dependent variable: HOME	Coeff.	Marginal effect
Male°	0.372*	0.092*
	(0.22)	(0.05)
City°	0.348*	0.085*
	(0.21)	(0.05)
Mark	-0.036**	-0.009**
	(0.02)	(0.004)
Age	0.458**	0.114**
	(0.21)	(0.05)
Age squared	-0.020**	-0.005**
	(0.01)	(0.002)
Business°	-0.788**	-0.184**
	(0.29)	(0.07)
Engineering°	-0.923**	-0.211**
	(0.43)	(0.09)
Law°	-0.233	-0.057
	(0.29)	(0.07)
University away°	-0.326	-0.080
	(0.22)	(0.05)
Master away°	-0.917**	-0.218**
	(0.29)	(0.06)
Internship°	-0.694**	-0.169**
	(0.26)	(0.06)
Internship at home°	1.030**	0.242**
	(0.32)	(0.07)
No_pref_HOME°	-0.997**	-0.225**
	(0.21)	(0.05)
Highmob°	-0.655*	-0.156**
	(0.34)	(0.07)
Unemployed°	3.614**	0.494**
	(0.46)	(0.07)
Masteryear	-0.225**	-0.056**
	(0.05)	(0.01)
Age*Highschool	-1.538**	-0.384**
	(0.56)	(0.14)
Age squared* Highschool	0.113**	0.028**
	(0.04)	(0.01)
Mark*Highschool	0.033	0.008
	(0.02)	(0.005)
Constant	4.177*	
	(2.28)	
N. Observations	700	
LL	-317.822	
Pseudo R2	0.330	

Standard errors in parentheses
* significant at 10%; ** significant at 5%
° Marginal effects for dummy variables are calculated as discrete change from 0 to 1

The probability of migration is increasing along the time span we are considering (*masteryear*). This finding is in line with the pattern of internal migration in Italy. Aggregate statistics show that, during the last five years, the number of people who migrated from the South to the North of Italy has increased considerably.

4. Where do migrants go?

4.1. Empirical Methodology

In section 3 we considered the individual migration decision in a binary framework, "move" versus "non-move". This approach allowed us to investigate the underlying differences between these two sub-groups of individuals, i.e. which individuals are more likely to migrate (or conversely to "stay home"). In this section we aim to provide a more in depth investigation of the regional characteristics affecting individuals' specific location decisions. The following analysis considers each individual's choice among the 20 Italian regions (including the home location).

The choice of the appropriate econometric methodology depends on assumptions regarding the decision-making process in terms of individual's migration. If migration is a sequential decision-making process, individuals will first make the decision whether or not to move and then (for the movers only) decide in which destination to relocate. Those behavioural assumptions underline the necessity to use a model which treats the decision to migrate separately from the choice between potential destinations. The natural candidate is a nested logit model or alternatively a model focusing on choice within a set of potential destinations only for the movers.

In a recent study, Davies, Greenwood and Li (2001) have argued that the decision to move and that of the destination choice cannot be considered as separated. Individuals jointly decide whether and where to move. We believe that this behavioural assumption is closer to reality. Methodologically this assumption on individual behaviour leads to a conditional logit model.[22]

The empirical model representing the migration decision might be derived from individual utility-maximising behaviour in the framework of a random utility model. Utility of an individual i at location j is given by:

$$U_{ij} = \beta' X + e_{ij} \tag{6}$$

where X is a vector of observable attributes pertaining to the potential destination regions and β is a conforming vector of utility coefficients, which might vary depending on characteristics of the decision-maker, and e_{ij} is a random disturbance. McFadden (1974) demonstrates that if the random disturbances in equation (6) are independent and identically distributed with the Weibull distribution, then the probability that an individual i will choose location j is given by:

$$P_{ij} = \frac{e^{(\beta' X_j)}}{\sum_{k=1}^{K} e^{(\beta' X_k)}} \tag{7}$$

---

[22] The conditional logit model can only identify choice-specific attributes (which might vary across individuals) but not individual-specific attributes or origin-specific attributes. Nevertheless, individual characteristics might be incorporated using interaction terms and a fixed-effect (a dummy variable equal to one if the potential alternative is the home region) can be used to capture non-observable factors explaining a greater attachment to the origin region.

where K is the number of alternative locations.[23]

Individual i chooses location j when U_{ij} is the maximum among the set K of potential destinations. The probability of locating in a particular region depends on its economic attractiveness but also on amenities and quality of life that the region is able to offer compared to other competing destinations.

In our analysis each individual i is paired with values of the independent variables (regional characteristics) relative to the year of completion of the course.[24] This approach has two positive effects. Firstly, the independent variables reflect regional characteristics at the moment when job-search activities were taken. Given the time span of our data, regional economic opportunities might have changed overtime. Secondly, we increase the variation in the independent variables, which is an advantage in the estimation of the conditional logit model.

Generally, the interpretation of the coefficients in logit models is not straightforward. One way of making this interpretation easier, in particular to assess their magnitude, is to express coefficients as average probability elasticities. As in Friedman *et al.* (1992), regional characteristics for each year are normalised to the national mean. Due to the normalisation, the mean of the independent variables is equal to unity, and the estimated coefficients have a convenient elasticity-like interpretation. The elasticity of the probability that an individual will choose region k with respect to an explanatory variable x_s can be calculated by differentiating equation (7):

$$E_{js}^i = \frac{\partial P_j}{\partial x_s} \frac{x_s}{P_j} = \beta_s (1 - P_j) \tag{8}$$

---

[23] We have performed tests for the Independence from Irrelevant Alternatives (IIA). The parameter estimates appear stable when we exclude sets of alternatives from the choice set.

[24] Due to lack of data this approach is not possible for a limited number of independent variables. In these cases we use the figures relative to the closest year available.

76

where P_j is the probability of selecting location j, x_s is the sth variable of the vector X, and β_s is the relative estimated coefficient.

If we sum across all individuals (i) and potential destinations (j) we obtain the relationship between average probability elasticity and the coefficient estimates, β:

$$\frac{E}{K} = \sum_{i=1}^{N} \sum_{j=1}^{K} E_{js}^{i} / K = \beta_s \frac{K-1}{K} \tag{9}$$

Since K the potential number of locations in our study is 20, elasticity can be easily obtained by multiplying the estimated β by 0.95. For the average region a one-percent increase in the values of independent variable x_s relative to the national mean (which is one) will cause a $\beta_s * (0.95)$ percent change in the estimated probability.

Normalisation has further advantages, it allows us (i) to remove the national growth trend for the independent variables over the period considered; (ii) to take into account the uneven pattern of growth across Italian regions; and (iii) to use similar scale for the independent variables used in the estimation and, therefore, to reduce computational problems.

4.2. Description of the data and variables influencing migration

The data on individual migration decision are derived from the postal survey as described in section 2. The conditional logit analysis is performed on migration data for 603 highly educated individuals that have benefited from the regional policy previously discussed. The sample under analysis is slightly different from the one used for the binomial logit, since we

exclude individuals who have yet to find employment and those who have migrated to a foreign destination. [25] Another difference with respect to the logit model is given by the fact that in this section we perform a conditional logit analysis which relates to the individual migration decision at the moment of the *first employment* following the completion of the postgraduate course financed by the regional authority.[26] This enables us to analyse, across individuals, a relatively homogeneous episode of personal decision making. Further relocations after the first employment can be influenced by a wide array of different reasons that it is difficult to control. However, it is notable that the difference in the spatial distribution of those individuals between the first and current employment is relatively small.

Figure 2 illustrates the regional distribution of the human capital "generated" by the local policy measure. Only 27.4% of these individuals are employed in Basilicata after the human capital investment period. A considerable share of individuals in the sample choose to work in Lombardia (mainly in the area of Milano) and Lazio. In this section we aim at a more in-depth investigation on the regional characteristics affecting individuals' location decision. For each individual in the sample, the data include 20 rows (one for each potential location).

The dependent variable **CHOICE** is equal to 1 for the region where the first job after the course was obtained and zero otherwise.

---

[25] We have decided to keep out of the sample individuals who have yet to find employment since they are mainly located in the home region. We leave out individuals who have selected a foreign destination due to practical difficulties in finding comparable destination specific data for the empirical modelling.
[26] The results discussed in section 3 looked at current employment (that is employment at the time when the survey was conducted).

Figure 2- The geography of the human capital generated: first employment
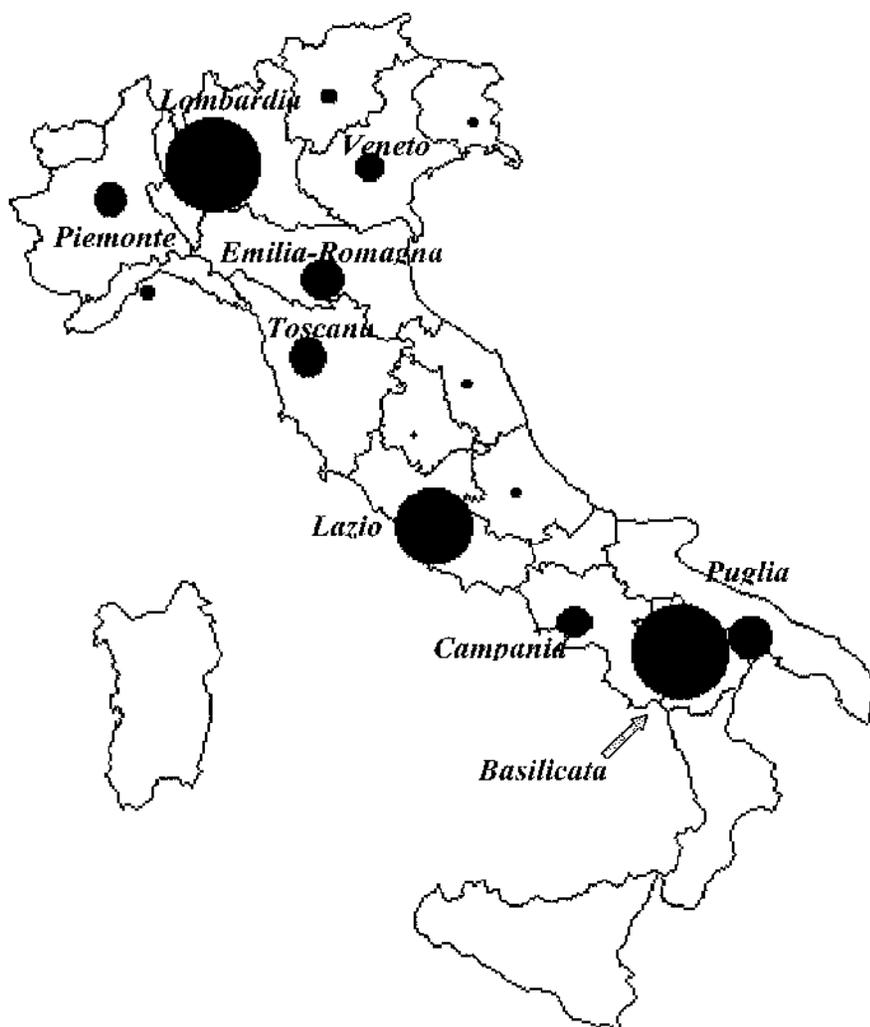
The explanatory variables used for the conditional logit are reported in *table 3*.

The variables thought to influence the migration decision include those reflecting economic conditions and quality of life along with some measures of individual migration costs associated with the potential destinations (based on the availability of information and the non-monetary psychological costs of moving).

As a measure of potential economic gains (or losses) from migration, we include in the model the regional *income per unit of labour*. A positive sign is expected for the coefficient on this variable, which means that individuals are attracted by locations with relatively higher income per unit of labour.

The regional *unemployment rate* is also used as a relative measure of job opportunities in a potential destination. We expect a negative sign on this variable since higher unemployment areas should deter immigration.

We use the GDP per capita growth rate in the 3 years preceding the end of the master course (*regional growth*) in order to test whether individuals are attracted toward regions outperforming the national growth trend. This variable is aimed at capturing the recent dynamics of the regional economy rather than the static aspect of economic success, which should be captured by the two measures described above. We also expect this variable to have a positive sign.

Of particular interest are those variables reflecting relative quality of life in potential destinations. Florida (2002) argues that location-specific characteristics such as lifestyle options, amenities and quality of life in general have an important function as catalyst for attracting and retaining talented individuals in U.S. cities. In a recent contribution Glaeser *et al.* (2000) conclude: "if cities are to remain strong, they must attract workers on the basis of quality of life as well as on the basis of higher wages". In our analysis we measure the impact of a set of quality of life variables on the probability that a certain region is selected by the individuals of our sample.

The variable *culture* represents expenditure per capita for fine arts and musical performances. This measure is a good proxy for demand, and therefore availability, of cultural amenities and is expected to positively influence the probability of migration to areas with a relatively higher value for this variable. We investigate also the effects of crime on the

migration decision by employing two variables. *Crime1* is the percentage of micro-crime offences out of the total number of declared offences.[27] These kinds of micro-offences are supposed to have a high impact on the day-by-day quality of citizens' life. Higher values relative to the national average are expected to have negative influence on migration. While this variable captures the "qualitative" aspects of criminal activity, it is not a good measure for its quantitative aspects. In this respect we use *crime2*, which measure the number of violent episodes of crime per 10,000 inhabitants.

In addition, we use an index of population concentration: *population density*. Unfortunately this variable is available only for year 1999, but we do not expect much variation in the time span considered here. This variable is measured as the ratio between the regional population living in the main city of each province within the region and the residual regional population living in smaller centres. High values of this measure imply more agglomeration into fewer cities. We do not have a priori expectation on the sign of this variable since it might capture both congestion costs and agglomeration forces.

---

[27] This index measures the number of offences like robberies in cars, apartments and shops, pick pocketing out of the total number of offences.

Table 3 – Modelling the location choice: description of explanatory variables employed in the conditional logit model

Variable	Description	Expected sign	Source
Income per unit of labour	Income per unit of labour (employee) (current Euros)	+	Istat
Unemployment rate	Regional unemployment rate (%)	-	Istat
Regional growth	Growth rate of the regional Gross Domestic Product (3 years)	+	Our calculations on ISTAT data
Preference Basilicata	Region specific dummy variable; equals 1 for Basilicata in the choice set if the individual have expressed a preference for the home region in her job-market search activity after completion of the postgraduate course,	+	Questionnaire
Distance	Distance from the province of origin (Potenza or Matera) to each Italian regional capital, Distance is expressed in 100 KM by car	-	Our calculations
Distance squared	Distance, as calculated above, squared	+	Our calculations
Distance adjusted	Distance from the province of origin as measured before, modified by assigning value "zero" to regions where the individuals have attended university, master studies or the internship	-	Our calculations
Distance adjusted squared	Distance adjusted, as calculated above, squared	+	Our calculations
Past migrants	Sum of migration outflows of individuals with high school or university degree from Basilicata to each Italian region, The variable is calculated summing flows for the last 5 years preceding the year of completion of the master course,	+	Our calculations on ISTAT data
Culture	Expenditure per capita at a regional level for theatrical and musical performances (thousand of lire - constant price 1999 - values have been deflated using the consumer price index)	+	Istat calculations on SIAE data
Crime1	Percentage of micro-crime offences out of the total number of declared offences	-	Istat
Crime2	Number of violent episodes of crime per 10000 inhabitants	-	Istat
Population density	Index of population concentration: resident population in the provincial capital / (total resident population in the province) - (resident population in the provincial capital)*100 (data 1999)	+/-	Istat
Undergraduate Studies Location	Dummy variable; equals 1 for the region where the individual attended university and 0 otherwise (note: University of Salerno and Bari are considered in this study "home" (Basilicata) locations due to geographical and cultural proximity)	+	Questionnaire
Postgraduate Studies Location	Dummy variable; equals 1 for the region where the individual attended postgraduate studies and 0 otherwise (note: Salerno and Bari are considered in this study home (Basilicata) locations)	+	Questionnaire
Internship Location	Dummy variable; equals 1 for the region where the individual attended the internship and 0 otherwise	+	Questionnaire

The migration literature suggests that the stock of established migrants originating from the same location is relevant in explaining the pattern of subsequent migration streams (see Carrington *et al.*, 1996; Greenwood, 1969). Past migrants reduce the cost of migration for other migrants by facilitating the flow of information on employment opportunities and reducing the costs of adaptation to a new environment by means of direct or indirect support. We test this hypothesis by using a variable *past migrants*, which is measured as the sum of past migrants from Basilicata to any potential destination over the last five years preceding the end of the master course.[28] We expect therefore that a location with a relatively large stock of immigrants from Basilicata (with respect to the national mean) to be a more attractive destination.

Distance is employed as a proxy for direct economic costs and psychic costs related to migration. This variable is measured in hundreds of kilometres by road between the main city of the province of origin and the capital city of each potential destination. Distance for the home region is zero. We use also *distance squared* in order to investigate non-linear effects eventually associated with increasing distance. As pointed out by Davies *et al.* (2001) a significant advantage of using the conditional logit for studying migration is the possibility to investigate the effect of distance. In models where the set of alternative choices is not explicitly considered, distance is simply ruled out. The adverse effect of distance on migration is generally attributed to two alternative explanations: (*i*) increasing psychic costs as individuals move from the region of origin (*ii*) the increasing costs and obstacles to information flows available to potential migrants on job opportunities in a set of destinations (Schwartz, 1973). These explanations should not be considered as mutually exclusive.

---

[28] The variable is created by summing flows of migrants with non-compulsory education levels (high school degree and above). As an alternative we have used a measure of total migration flows and even if the qualitative nature of the results does not change, the variable used in the analysis has a stronger explanatory power. Since it is difficult to assign a value for this variable in correspondence to the home location (Basilicata), we decided to consider the effects of the variable only for the migrants. This approach is expected to control for any bias in the estimation of the relative coefficient.

Since we are studying the migration decision of individuals coming from the same region of origin, we obviously observe in the data a higher propensity to select the home region. Some individuals will have an attachment toward this region which is explained by a set of economic and non-economic unobservable factors. The inclusion of a dummy variable for Basilicata (home-region fixed effect) should capture the effects of these unobservables. Nevertheless, individuals might or might not have a "home" attachment; this is particularly important given the nature of our sample, people who are likely to have a looser attachment to their home region compared with the rest of the population. In the questionnaire we have asked individuals about their geographical preferences in the job-market search after completion of the master course. We use this information to generate a dummy variable (*Preference Basilicata*) which is equal to one for region Basilicata in the choice set only for those individuals who have expressed a geographical preference for the home region. This should capture the "home" region attachment only for the individuals who really have such attachment.

4.3. Discussion of the Empirical Results

The results of the conditional logit estimations are presented in *table 4*. As mentioned in the methodology section, since the explanatory variables have all been normalised (apart from distance variables) relative to the national average, the magnitude of the coefficient corresponds to its relative importance. Coefficients on non-dummy variables have an elasticity-like interpretation as described in equation (9).

Overall, the explanatory variables are statistically significant and have the expected sign. *Version (A)* presents estimation of a quite parsimonious model with a specification that is similar to others employed in migration literature. All variables are significant and have the

expected influences on the attractiveness of a region. The second column, *version (B)*, shows the contribution to the model estimation given by the inclusion of quality of life measures. The importance of quality of life variables for migration decision of highly educated and skilled individuals is confirmed by the fact that the variables *culture, crime1* and *crime2* are highly significative. Individuals are attracted towards location with higher cultural amenities and deterred from locations with diffused micro-criminality.

Version (C) and *version (D)* of the model display the importance of taking into consideration the effect of a "subjective" rather than a physical measure of distance. Migration is an important and complex decision for individuals and the amount of information required to make a choice might be quite substantial. A potential migrant will form his/her own subjective predictions on future income (or more generally, utility) streams in a given destination. Even within a rather homogeneous group of individuals as in this study, the available information and personal connections will differ greatly. The probability that an individual will select a given location will increase the larger is the amount of information available on job opportunities in that location. It is an acknowledged fact in migration studies that the quantity and quality of information are negatively affected by distance (Lucas 2001). Knowledge about the destination region depends upon personal contacts and upon sources of information which are not universally available. Physical distance is a weak approximation for the difficulties in accessing the needed information.

Instead of using physical distance as a proxy for these costs, thanks to the quality of our data we are able to use a more appropriate measure of relative distance, which takes into account individuals' past migration experiences. The hypothesis under analysis here is that individuals with previous migration experience of a particular destination have the opportunity to acquire information on the local labour market at low costs. It is also likely that non-monetary psychological costs of migration will be lowered as a result of increased

familiarity with the local environment. It is to be expected that these individuals will have a higher probability of selecting that location. Our model uses information on individuals' past migration experience during the undergraduate, and postgraduate study periods and during any period spent in an internship.

In *version (C)* we use two new variables *distance adjusted* and its square (see *table 3* for explanation) which incorporates information on the individual past migration experiences. As we can see, these two variables are highly significative and substantially improve the overall performance of the model. The estimated coefficients of these two variables do not differ dramatically from the previous two adopted. By using this specification we implicitly assume that past migration during the university studies, master studies and those for the internship have the same effect on the probability of selecting a certain location. Of course this might not be true since they represent episodes of an individual's life with a different proximity to individual job-market search activities. For instance we might expect that the period of the internship and master course was more job-oriented than the university period. The importance of prior migration experiences is captured in three separate dummy variables in *version (D)*: *Undergraduate Studies Location, Postgraduate Studies Location and Internship Location* (see *table 3* for details). The internship considered here is part of the postgraduate course financed by the regional authority and is not necessarily done in the same region where the course was held. As expected, and not surprisingly, an internship in a given location substantially improves the probability that the individual will remain there. This effect is stronger compared to that of *Undergraduate* and *Postgraduate Studies Location*, which are nevertheless quite important. The importance of these dummy variables in our study underlines how relevant is the availability of information on potential destinations for the individual's migration decision process. A lack or poor quality of information on local

economic opportunities represents a market failure that can be, at least partly, corrected through the use of simple policy measures.

Turning to the rest of coefficient estimates, the model confirms that economic opportunities in the potential destination are relevant for the location selection process. In *version (D)* a one-percent increase in income per unit of labour relative to the national mean has the strongest effect on migration probability (5.76% increase).[29] Regions with higher unemployment rates deter migration. This result is in line with previous migration studies that have used micro-data and confirms the hypothesis that the unemployed are particularly sensitive to local unemployment rates.[30] Faster-growing areas experience an above average immigration of highly skilled individuals as indicated by the positive and significant coefficient on regional GDP growth. This last effect, according to the estimated parameter, has a rather small magnitude.

The dummy variable for Basilicata is highly significant and positive. As mentioned before, this variable captures the effects of those unobservable characteristics ("home" attachment) for the individuals who have expressed a geographical preference for Basilicata.

We find that the migration probability declines with the distance between the origin and potential destination. There is strong support to conclude that this effect is non-linear. The negative effect of distance declines at greater distances (the marginal impact on the probability of migration toward a potential destination is decreasing in distance).

The existing stock of highly educated migrants from the region of origin *(past migrants)* plays an important role in explaining which location is selected among alternatives. This variable should capture a lower cost of migration towards the potential destination through ease of access to information and direct and indirect support. But it may

---

[29] For a hypothetical region with average characteristics in year 2000, this means that an increase in income per unit of labour of Euros 280 (national average was Euros 28.050, in current prices) will lead, ceteris paribus, to a 5.76% increase in the probability of attracting highly educated individuals.

also capture the quality and quantity of economic opportunities in the destination region since the stocks of past migrants are generally proportional to the size of the population in alternative locations. In migration studies, it is common for the size of regional population or the level of GDP to be used as control variables capturing the quantity of opportunities. A larger region has a larger pool of locations to select and therefore potentially more opportunities. In our study these control variables were found not significant. Perhaps this result reflects the fact that the Italian Mezzogiorno is composed of some densely populated regions (such as Sicilia) for which it is hard to believe that population is a good proxy of economic opportunities available to the potential migrants.

---

[30] Frequently migration studies based on aggregate data show mixed results on the sign and significance of the unemployment rate variable.

Table 4 – Modelling the locational choice: conditional logit estimates

Dependent variable: CHOICE	Mod (A)	Mod (B)	Mod (C)	Mod (D)
Income per unit of labour	18.164** (0.99)	13.881** (1.41)	7.168** (1.42)	6.060** (1.74)
Unemployment rate	-1.382** (0.28)	-0.767* (0.38)	-0.692* (0.36)	-0.752* (0.45)
Regional growth	0.623** (0.09)	0.299** (0.10)	0.465** (0.11)	0.368** (0.12)
Preference Basilicata	1.915** (0.25)	1.583** (0.26)	1.849** (0.24)	1.661** (0.31)
Distance	-1.088** (0.11)	-1.415** (0.16)		-0.909** (0.19)
Distance squared	0.061** (0.01)	0.098** (0.01)		0.071** (0.01)
Past migrants	1.508** (0.07)	1.675** (0.09)	1.583** (0.09)	1.567** (0.10)
Culture		1.954** (0.35)	1.057** (0.34)	1.436** (0.41)
Crime1		-4.864** (0.74)	-5.535** (0.84)	-5.078** (0.91)
Crime2		-0.776** (0.31)	-0.597* (0.31)	-0.800** (0.36)
Population density		0.083 (0.15)	0.289* (0.14)	0.292 (0.18)
Distance Adjusted			-1.129** (0.09)	
Distance Adjusted squared			0.086** (0.01)	
Internship location				2.316** (0.14)
Undergraduate studies location				0.966** (0.15)
Postgraduate studies location				0.9795** (0.14)
LL	-1018.88	-985.02	-707.82	-652.99
McFadden's R2	.436	.455	.608	.639
BIC	-1777.79	-1819.90	-2374.30	-2464.74
Standard errors in parentheses * significant at 10%; ** significant at 5% N = 12060				

89

Table 5 - Conditional logit: interaction effects

Variables	Version (E)
Income per Unit of Labour	6.207**
	(1.76)
Unemployment rate	-2.139**
	(0.75)
Regional growth	0.40**
	(0.12)
Preference Basilicata	1.735**
	(0.31)
Distance	-0.790**
	(0.22)
Distance squared	0.058**
	(0.02)
Past migrants	1.636**
	(0.11)
Culture	1.467**
	(0.42)
Crime1	-5.421**
	(0.93)
Crime2	-0.890**
	(0.37)
Population density	0.303*
	(0.18)
Internship location	2.110**
	(0.16)
Undergraduate studies location	0.592**
	(0.25)
Postgraduate studies location	1.237**
	(0.16)
City * Distance	-0.242*
	(0.14)
City * Distance squared	0.027*
	(0.01)
City * Undergraduate studies location	0.595**
	(0.30)
Mark * Unemployment rate	0.033**
	(0.01)
Master law * Unemployment rate	0.585**
	(0.29)
Master law * Postgraduate studies location	-0.839**
	(0.31)
Master law * Internship location	0.884**
	(0.35)
Master engineering * Undergraduate studies location	1.193**
	(0.44)
Master arts * Postgraduate studies location	1.257**
	(0.58)
LL	-629.86
McFadden's R2	0.651
Standard errors in parentheses * significant at 10%; ** significant at 5%; N,=12060; BIC=-2453,30	

Along with economic opportunities, *quality of life* is a very important aspect that should be taken into account by regional policymakers who aim at making their locations attractive to talented individuals. The weight attached to these amenities is considerable. Crime is a strong push factor. This is particularly true for highly educated individuals who are generally more footloose and therefore able to "react" more easily to such disamenities by relocating to safer areas. In our model the combined negative effects of *crime1* and *crime2* are remarkably high. Diffused criminality discourages the potential location decision by talented individuals and investors. Finally, population density has a weak positive effect on migration probability suggesting that highly skilled individuals are attracted toward more urbanised areas where agglomeration forces are strong and it is easier to have access to a more variegated and sophisticated bundle of goods and services. This result should be taken with caution since the variable appears insignificant in some versions of the model.

4.4 The relevance of personal characteristics: some interaction effects

In the conditional logit the effect of personal characteristics on migration cannot be directly investigated. Instead we need to interact these variables with place characteristics. The results of the analysis of the importance of personal characteristics are presented in *table 5*.

Interactions between the dummy variable *city* with *distance, distance square* and *undergraduate studies location* are found significant. As we have seen in the logit model, individuals who live in a larger agglomeration have a lower probability to migrate. The negative effect of distance for those born and living in a city is stronger and significant at the 10% level. A potential destination which is 100 km further away from a region with average distance from Basilicata (520 km circa) has a probability of being selected which is −0.56% lower for those born in a city and −0.46% lower for those who are not. The interaction

between *city* and *undergraduate studies location* has a positive sign, significant at 10% level. For these individuals the probability of remaining in the same region where university was attended is higher. With regard to this, we speculate that the choice of university location for these individuals already probably discounts a certain amount of inertia.[31]

The interaction between the mark obtained at the end of bachelor's degree and the unemployment rate (*mark*unemployment rate*) is positive and significant. This result indicates that the most able individuals among the sample are deterred less from areas with a higher unemployment rate. This is not surprising since these individuals are more able to compete in the job-market even in local economies with a reduced set of opportunities.[32] Finally we study the effects of the subject of master studies on migration probabilities. Individuals with a post-graduate education in law and political science are less discouraged by high unemployment area (*master law* unemployment rate*). A considerable share of these individuals will continue their training to become lawyers or fiscal advisors after the university and master. In Italy, for these professions a training period of two to three years after the bachelor's degree is required. For this reason a apprenticeship period is compulsory for embarking on these careers. Generally it is easier to obtain it in a location where the individual has strong personal connections. This explains the positive interaction of *master law* (master in law and political science) with the dummy variable for *undergraduate studies* and *internship location*, and the negative interaction with the dummy on *postgraduate studies' location*. The period of time spent during the master is probably not enough to build the necessary connections to obtain a apprenticeship position.

Individuals with a degree in engineering or architecture (*master engineering*) are more likely to locate in the region where their university course was held, while the region

---

[31] A tabulation of the two variables confirms that individuals born in Matera and Potenza are slightly more likely to study in a "home" university.

[32] The introduction of this variable significantly reduces the standard error estimate for the unemployment rate measure in the model.

where the master was attended is more important for individuals with education in literature studies, languages and arts in general (*master arts*). We also found weak support for the fact that individuals with a more business oriented educational background are more likely to migrate toward more densely populated areas.

It is worth mentioning that the migration behaviour with respect to gender does not show any significant difference in the conditional logit. The interaction effects of the dummy variable *male* were all statistically insignificant.

5. Final comments

The existing migration literature points out that the propensity of highly skilled and educated individuals to migrate is high when compared to the rest of the population (Antolin and Bover, 1997; Haapanen, 2003; Hughes and McCormick, 1989; Molho, 1987; Levine, 1996; Ritsila and Ovaskainen, 2001). However, to the best of our knowledge, few studies have conducted an in depth investigation into the factors explaining the geographical pattern of their decisions.

In this paper we focus our attention on the micro-level location decisions of highly skilled and educated individuals, who constitute one of the most important resources in modern economies. We employ data on a sample of individuals residing in a small peripheral Italian region (Basilicata) who have benefited from a locally funded human capital investment policy. Our results show that even within a quite homogeneous group of individuals, there are key personal characteristics that affect the migration propensity.

Understanding the factors explaining the location decision of talented individuals is of considerable importance, particularly in the light of increasing regional economic integration which makes this group of individuals extremely footloose.

As highlighted in this paper, among the factors explaining individuals' location decision, quality of life is very important. For this reason policy interventions aimed at improving the quality of citizens' lives should be regarded as very attractive. Diffused criminality discourages potential location decisions by talented individuals and investors. We believe this is a quite important factor explaining the poor economic performance for some areas of the Italian Mezzogiorno, which otherwise should be considered particularly attractive given the high quality and large quantity of human capital available. This consideration is not

intended to discount the importance of regional economic performance, which should be considered as a pre-requisite for attracting highly skilled and talented individuals. We want to emphasise that improving quality of life can lead to a substantial pay-off in terms of future growth and prosperity of peripheral regions.

Finally, we are able to assess the relative importance of the availability of information on potential destinations for the individuals' migration decision process. Lack or poor quality of information on local economic opportunities represents a market failure that can be, at least partly, corrected by adequate policy measures. Actions aimed at this purpose might, therefore, be considered as strategic in trying to maintain and attract human capital.

References

Antolin, P. and O. Bover, 1997, Regional migration in Spain: the effect of personal characteristics and of unemployment, wage and house price differentials using pooled cross-section, *Oxford Bulletin of Economics and Statistics* 59, 215-235.

Borjas, G., 1987, Self-Selection and the Earning of Immigrants, *American Economic Review* 77, 531-553.

Borjas, G., S. G. Bronars and S. J. Trejo, 1992a, Assimilation and the earnings of young internal immigrants, *Review of Economics and Statistics* 74, 170-175.

Borjas, G., S. G. Bronars and S. J. Trejo, 1992b, Self-Selection and internal migration in the United States, *Journal of Urban Economics* 32, 159-185.

Carrington, W., E. Detragiache and T. Vishwanath, 1996, Migration with Endogenous Moving Costs, *American Economic Review* 86 (4), 909-930.

Coniglio, N. D., 2003, Regional Integration and Migration: an Economic Geography Model with Heterogeneous Labour Force, Glasgow University Department of Economics Discussion Paper series, n.1/2003.

DaVanzo, J., 1978, Does unemployment affect migration? Evidence form microdata, *Review of Economics and Statistics* 60, 504-514.

DaVanzo, J., 1983, Repeat migration in the United States: who moves back and who moves on?, *Review of Economics and Statistics* 65, 552-559.

Davies, P., Greenwood M. and H. Li, 2001, A Conditional Logit approach to U.S. State-To-State Migration, *Journal of Regional Science* 41, 337-360.

Farber, S. C., 1978, A directional flow migration model, *Southern Economic Journal* 45, 205-217.

Florida, R., 2002, Bohemia and Economic Geography, *Journal of Economic Geography* 2 (1), 55-71.

Friedman, J., Gerlowski D. and Silberman J., 1992, What Attracts Foreign Multinational Corporations? Evidence from Branch Plant Location in the United States, *Journal of Regional Science* 32, 403-418.

Gathak, S., Levine P. and Price S. W., 1996, Migration theories and evidence: an assessment, *Journal of Economic Surveys* 10 (2), 159-198.

Glaeser, E. L., Kolko J. and Saiz A., 2000, Consumer City, NBER Working Paper Series, n. 7790, July.

Greenwood, M. J., 1969, An Analysis of the Determinants of Geographic Labor Mobility in the United States, *Review of Economics and Statistics* 51 (2), 189-194.

Haapanen, M., 2003, Studies on the Determinants of Migration and the Spatial Concentration of Labour, PhD thesis, Jyvaskyla: University of Jyvaskyla.

Herzog, H. W. and Schlottmann A. M., 1981, Labour force migration and allocative efficiency in the United States: the role of information and psychic costs, *Economic Inquiry* 19, 459-475.

Herzog, H. W. and Schlottmann A. M., 1984, Labour force mobility in the United States: migration, unemployment, and remigration, *International Regional Science Review* 9, 43-58.

Herzog, H. W., Schlottmann A. M., Boehm T. P., 1993, Migration as spatial job-search: a survey of empirical findings, *Regional Studies* 27 (4), 327-40.

Hughes, G. A. K. and McCormick B., 1989, Does migration reduces differential in regional unemployment rates?, in: J. Van Dijk, H. Folmer, H. W. Herzog and A. M. Schlottmann, eds., *Migration and labour market adjustment* (Kluwer, Dirdrecht).

Ledent, J., 1990, Canada, in: C. B. Nam, W. J. Serow and D. F. Sly, eds., International handbook on internal migration (Greenwood Press, New York).

Levine, P., 1996, Migration Theories and Evidence: an Assessment, *Journal of Economic Surveys* 10, 159-98.

Long, L., 1992, Changing residence: comparative perspectives on its relationship to age, sex, and marital status, *Population Studies* 46, 141-158.

Justman, M., Thisse J. F., 1997, Implications of the mobility of skilled labor for local public funding of higher education, *Economic letters* 55, 409-412.

Krugman, P., 1991, Increasing Returns and Economic Geography, *Journal of Political Economy* 99, 483-499.

Lucas, R. E., 1988, On the Mechanics of Economic Development, *Journal of Monetary Economics* 22, 3-42.

Lucas, R. E.B., 2001, The Effects of Proximity and Transportation on Developing Country Population Migration, *Journal of Economic Geography* 1, 323-339.

McFadden, D., 1973, Conditional Logit Analysis of Qualitative Choice Behaviour, in: P. Zarembka, eds., *Frontiers in Econometrics* (Academic Press, New York).

McHone, W. W. and Rungeling B., 1993, Replacement income, duration of unemployment and the migration, *Applied Economics* 25, 905-909.

Molho, I., 1987, The migration decisions of young men in Great Britain, *Applied Economics* 19, 221-243.

Molle, W. T. M. and van Mourik A., 1989, A static explanatory model of international labour migration to and in Western Europe, in: Gordon I. and Thirwall A. P., eds., *European factor mobility: trends and consequences*, (Macmillan, London).

Otomo, A., 1990, Japan, in: C. B. Nam, W. J. Serow and D. F. Sly, eds., *International handbook on internal migration*, (Greenwood Press, New York).

Pissarides, C. A. and Wadsworth J., 1989, Unemployment and the inter-regional mobility of labour, *Economic Journal* 99, 739-755.

Plane, D. A., 1993, Demographic Influences on Migration, *Regional Studies* 27, 375-383.

Ritsila, J. and Ovaskainen M., 2001, Migration and regional centralization of human capital, *Applied Economics* 33, 317-325.

Robinson, C. and Tomes N., 1982, Self Selection and Interprovincial Migration in Canada, *Canadian Journal of Economics* 15, 474-502.

Schultz, T. W., 1962, Reflections on Investment in Man, *Journal of Political Economy* 70 (2), 1-8.

Schwartz, A., 1973, Interpreting the Effect of Distance on Migration, *Journal of Political Economy* 81 (5), 1153-1169.

Temple, J., 2001, Growth effects of Education and Social Capital in the OECD Countries, *OECD Economic Studies* 33, 57-101.

Chapter 4

Conduct of the survey, design of the questionnaire and sample selection issues

Nicola D. Coniglio

1. Introduction

The data employed in the analysis have been collected directly by the author by means of a postal survey. A comprehensive assessment of the research presented in the previous chapter cannot leave aside the methodological procedure used for generating the data set. In addition, the methodological process in itself can be considered a worthwhile contribution, given the possibility to replicate it in similar, future research. The aim of this chapter is to provide details of the data collection method that has been employed, outline the different steps of the survey process and describe the design of the questionnaire. Moreover, we discuss sample selection issues in our empirical investigation.

This chapter is organised as follows. In section 2, we define the research issues which constitute the root of the survey and the associated data requirement. In section 3, we motivate our choice of adopting a postal survey and present the initial phases of the questionnaire's design. The conduct of the survey and the rate of response are outlined in section 4. Section 5 contains a digression on the sample selection issue. The covering letter and the questionnaire are reported in the appendix.

2. Research issues and data requirement

The definition of the research objectives and a clear investigation of the associated data requirements are the backbone of any rigorous empirical investigation. The first step of our survey was devoted to determining the main questions we wanted to address. As highlighted in previous chapter, our focus is on the micro-level migratory behaviour of a particularly important stratum of the population: highly talented and skilled individuals. Knowledge of

the factors explaining the location decision of talented individuals is of considerable interest, in particular in the light of increasing regional economic integration which makes this group of individuals extremely footloose. If policy makers, in particular in peripheral regions, are to capitalise on highly skilled and educated individuals, regarded as one of the most important resources in modern economies, a first step is to understand the migratory behaviour of such individuals. This knowledge can then be used in a strategic way to formulate talent attraction and retention policies.

The individuals who benefited from the human capital investment policy adopted by Regione Basilicata constitute a natural experiment for making inference about this particular stratum of the population. The regional authority imposes no requirements on the subject and location of the course financed. The selection criteria concern educational attainment of the candidate and the "quality" and reputation of the educational institution which offers the courses.[1] The policy is very selective toward high quality candidates representative of the population of young and highly educated individuals to whom our attention is focused.

Our analysis explores the factors affecting the location decision of highly educated individuals in an economy possessing all the features of a core-periphery structure.[2] In the previous chapter the empirical analysis aimed at addressing the following two sets of questions:

1) Who are the migrants? Is there a significant difference according to some personal characteristics between migrants and non-migrants?

---

[1] These regulations are aimed at avoiding courses of questionable quality. An applicant with a successful educational record willing to embark upon further investment in their education will very likely obtain regional support. The rejection rate is very small and is based mainly on considerations that pertain to qualitative aspects of course choice.

[2] Of course, as with all case studies, it is necessary to exercise caution when generalising the findings to other situations. Nevertheless, it is reasonable to argue that the behaviour of the highly educated and skilled individuals in our sample is not substantively different from that of similar individuals in peripheral regions of many developed countries.

2) Where do migrants go? What are the regional characteristics that make a region attractive for highly educated and talented individuals?

In addition to the outlined main research issues, other objectives were defined during the initial phase of survey design. These include investigations concerning the labour market performance of the sample under analysis (duration of the unemployment spells, use of different job search methods, etc.). Another set of research issues, influential in the design of the survey, relates to the specific policy undertaken and is mainly aimed at providing Regione Basilicata with useful information for assessing and improving the policy measure adopted (quality of the course financed, importance of the regional contribution, geographical diffusion of the policy etc.). Subsequent phases of survey design, in particular the definition of the items to be included in the questionnaire, reflect these planned research objectives.

Diagram 1 gives a stylised picture of the individual data collected in order to address the stated research questions. In particular, we collected data on: (i) personal characteristics which might affect migratory behaviour (sex, age, level and type of education etc.); (ii) the course financed by the regional authority (quality, location, duration etc.); (iii) job market search activity (methods used to find a job, preferred location, unemployment spell); and (iv) space-time career details from first to current employment status.

Once the data requirement for the empirical investigation was established the second phase implied the definition of the data collection method and the design of the questionnaire.

Diagram 1 - Data requirement: a stylised picture

```
┌──────────────────┐        ┌──────────────────┐
│    Individual     │   ═▷   │  Higher education │
│  characteristics  │        │   (undergraduate  │
│                   │        │      studies)     │
└──────────────────┘        └──────────────────┘
                                     │
                                     ▽              ┌──────────────────┐
                                                    │      Policy       │
                                              ◁═════│   Intervention    │
                                                    └──────────────────┘
                            ┌──────────────────┐
                            │                   │
                            │ Further education │
                            │                   │
                            └──────────────────┘
                                     │
                                     ▽              ┌──────────────────┐
                                                    │    Job-search     │
                                              ◁═════│     strategy      │
                                                    └──────────────────┘
                            ┌──────────────────┐
                            │                   │
                            │    Employment     │
                            │                   │
                            └──────────────────┘
                            ↙                   ↘
              ┌──────────────────┐    ┌──────────────────┐
              │       Home       │    │       Away       │
              └──────────────────┘    └──────────────────┘
```
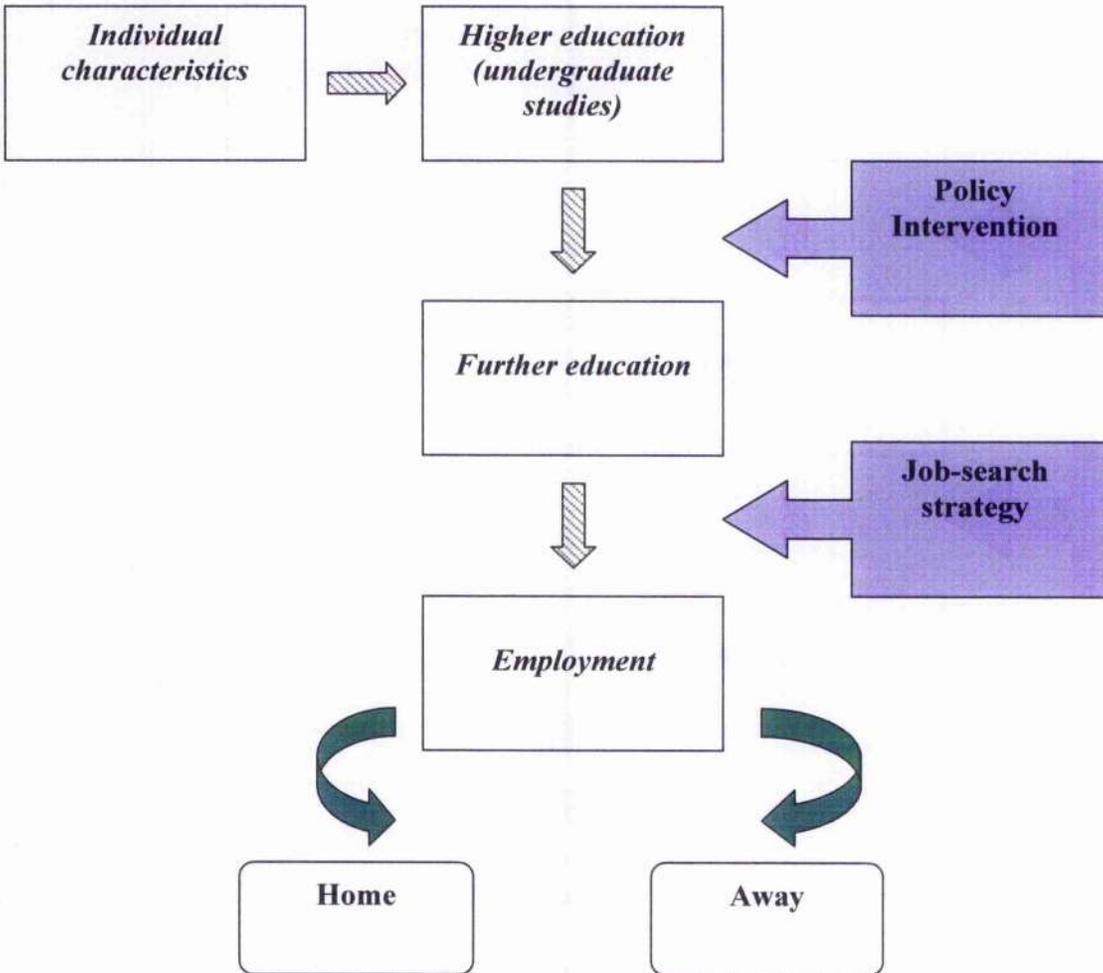
3. Data collection method and design of the questionnaire

Given the relatively limited number of beneficiaries of the regional policy intervention, i.e. slightly more than one thousand, the survey was administered to all of them. For this reason, sampling was not an issue in our survey.[3]

For reasons mainly related to the cost of the survey, the data collection method chosen was a self-administered questionnaire mailed to each individual address. Different administration modes typically have different non-response rates, the highest being generally in postal surveys. The main disadvantage of postal surveys is associated with the "response inertia" which might seriously reduce the percentage of returned questionnaires and, therefore, might compromise the empirical analysis. The main factors affecting this inertia are the following:

1) length of the questionnaire;

2) cost of returning the questionnaire;

3) distrust by the respondent of the aims of the survey;

4) 'sensitive items', for instance those regarding income, which might induce respondents not to return the questionnaire or to leave it incomplete.

Particular attention was devoted to minimising the response inertia.

Keeping in mind the data requirements, a first version of the survey questionnaire was constructed using the following key principles in designing each question:

- The respondent should clearly understand what she or he is being asked.

- The respondent can be expected to know the answer to the question.

- The respondent should easily understand how the answer should be given.

---

[3] Generally, the sampling strategy is one of the initial steps of survey design. It is a very delicate issue since the reliability of a survey and also its costs increase as the sample size increases.

After this step was taken, the questionnaire was administered to 15 individuals who had attended a postgraduate course (not necessarily financed by Regione Basilicata, hence not part of our final sample). Selected individuals were asked to complete a preliminary version of the questionnaire, which met all the requirements pointed out in the previous section. After completion of the submitted questionnaires, we investigated respondents' opinions about the length of the questionnaire and the complexity and possible ambiguity of each single item (**pre-test**). Several modifications of the questionnaire were undertaken as a consequence of the pre-test. In addition, sensitive items and other items that generated some confusion to the respondent were identified.

Two competing formats of the questionnaire were defined: *version (A)* and *version (B)*. Both versions were designed to gather the same set of information from the respondent, but some of the items were formulated in different ways. For instance, the item regarding the current gross yearly income was designed as 'open question' in version (A) and as 'closed – ranked question' in version (B). We conducted a **split panel test**, with the two formats, in order to evaluate the impact of wording and question order on the rate of response. In addition, we checked for the complexity or ambiguity of some items not revealed by the pre-test. Forty individuals were randomly selected from the list of beneficiaries of the regional policy. We submitted to each half of the sample a different version of the questionnaire.

Table 1: Questionnaire design: main results of the split panel test

	Number of questionnaires sent	Number of questionnaires returned	Rate of response	Average time of response (days)
Version A	25	10	40%	45
Version B	25	14	56%	40

The split panel test took place between February and April 2002. On the basis of the evidence of the test (see **table 1**), we decided to select version (B) of the questionnaire. In fact, version (B) performed better than version (A) both in terms of rate of response (56% versus 40%, respectively) and, to a lesser extent, in terms of average time of response (40 versus 45 days). Further useful information was collected on possible improvements of the questionnaire during this phase.

Using pre-stamped freepost envelopes solved the issue concerning the cost to the respondent of returning the questionnaire. In order to increase the motivation to fill in the questionnaire, in a cover letter: (i) we informed the respondents on why the survey was being conducted; (ii) we stressed the importance of their answer as a means of improving the effectiveness of the policy; (iii) and reassured them about the confidentiality of the data collected (see **appendix 1**). The name of the designated respondent was unequivocally indicated in the questionnaire and in a cover letter. All the postal correspondence and inquiries were directed to an official of the regional authority as a guarantee of the reliability of the survey.

4. Conduct of the survey and rate of response

The survey questionnaire, resulting from the phases outlined above, was mailed in May 2002 to all the individuals who benefited from the local policy measure from 1991 (the first year in which this measure was implemented) to 2001, which amounts to slightly more than one thousand individuals.[4] The English version of the questionnaire is reported in **appendix 2**.

After 45 days from the initial dispatch, a letter of reminder and a new copy of the questionnaire was mailed to non-respondents. In order to reduce 'response inertia', these individuals were subsequently contacted by phone. In case the individuals had changed residence, the new address was collected and a new questionnaire was sent.

By mid-September, a total of 740 individuals returned the questionnaire, approximately 70% of the total. After deleting observations relating to incomplete questionnaires, the usable sample contains 700 observations (see **table 2**).

Table 2 – Description of non-usable questionnaires

Motivation	Number
Missing relevant information on undergraduate studies (location, final mark or year of completion)	14
Missing relevant information on career history (only sections (a) and (b) of the questionnaire were completed)	32
Returned blank	5

Note: the total from this table differs from 40 since more than one motivation for non-usability might co-exist for the same questionnaire.

---

[4] Approximately 3.5% of the questionnaires were returned unanswered because of wrong address provided to us by the regional authority.

5. *Sample selection issue*

Sample selection problems arise when the sample studied represents a sub-group which is not representative of the entire population whose behaviour is under analysis. Given the nature of our sample, young and highly educated individuals, our results are intended to be generalised to the fraction of the population composed only by young, talented and skilled individuals in peripheral regional economies such Basilicata. The most likely causes of sample selection problems in micro-data based migration studies are the following (see also Greenwood, 1997): (1) *sampling design and population coverage*; (2) *time-dependent disturbances* and (3) *different rates of response to the survey*. We briefly comment, using the same order, on the relevance of these problems to the data employed in our work. With regard to the first point, as previously said, our analysis is conducted on highly educated individuals who have benefited from a measure aimed at financing individuals' human capital formation. These individuals represent the upper tail of the regional population educational attainment distribution. In principle this sample is a good proxy for the targeted population we aim to analyse.[5]

The decision tree for an individual wishing to obtain policy support is illustrated in **diagram 2**. In stage 1, highly educated individuals can choose to apply to the regional authority in order to get financial support for further investment in their human capital accumulation. Since the regional authority imposes no requirements on the subject and location of the course that could be attended and given the high public knowledge about this policy intervention, the potential problem of sample selection at this stage is rather small. Presumably those individuals with very high costs associated with a further investment in human capital will not apply, but this is exactly what we want, namely to include only

talented and skilled individuals. In stage 2, an individual who has applied might be accorded funds or not. The selection criteria imposed by the regional authority concern educational attainment of the candidate and the "quality" and reputation of the educational institution which offers the courses. Once again, the policy is selective towards high quality candidates. We therefore argue that the sample under scrutiny (right-hand-side of the decisional tree at stage 3) is representative of the population of young and highly educated individuals at whom our attention is focused.

The problem of time dependent disturbances is probably not very serious since we consider individual migration decision over a decade. Even if changes in the general pattern of migration are possible, we do not expect them to be dramatic. In addition, in the second part of our investigation (conditional logit model) we avoid the problem by using time-dependent explanatory variables.

With regard to the third point, the questionnaire was sent to all the beneficiaries of the policy from its start in 1991 until 2001. This amounts to approximately one thousand individuals. Since the rate of response to our postal survey was considerable (70%), we believe that a strong selection bias due to non-responses is very unlikely.[6]

---

[5] The number of observations at our disposal is considerable if compared to the total regional population that in the last census was slightly less than six hundred thousands (individuals between 25-29 years were 48 thousand in 1997).

[6] An investigation into the causes of non-response was conducted by tracking a number of non-respondents. Apart from the usual inertia in answering to postal surveys, a relevant share of the non-responses was due to inaccurate transcription of individual postal addresses in the documentation provided by the regional authority.

Diagram 2
Potential sample selection: individual's decision tree

STAGE 1:

APPLY FOR
THE REGIONAL
FUNDS OR NOT

STAGE 2:
IF NOT:

MIGRATE OR
NOT

STAGE 2:
IF YES:

FUNDED OR
NOT FUNDED

STAGE 3:
IF NOT:

MIGRATE OR
NOT

STAGE 3:
IF YES:

MIGRATE OR
NOT
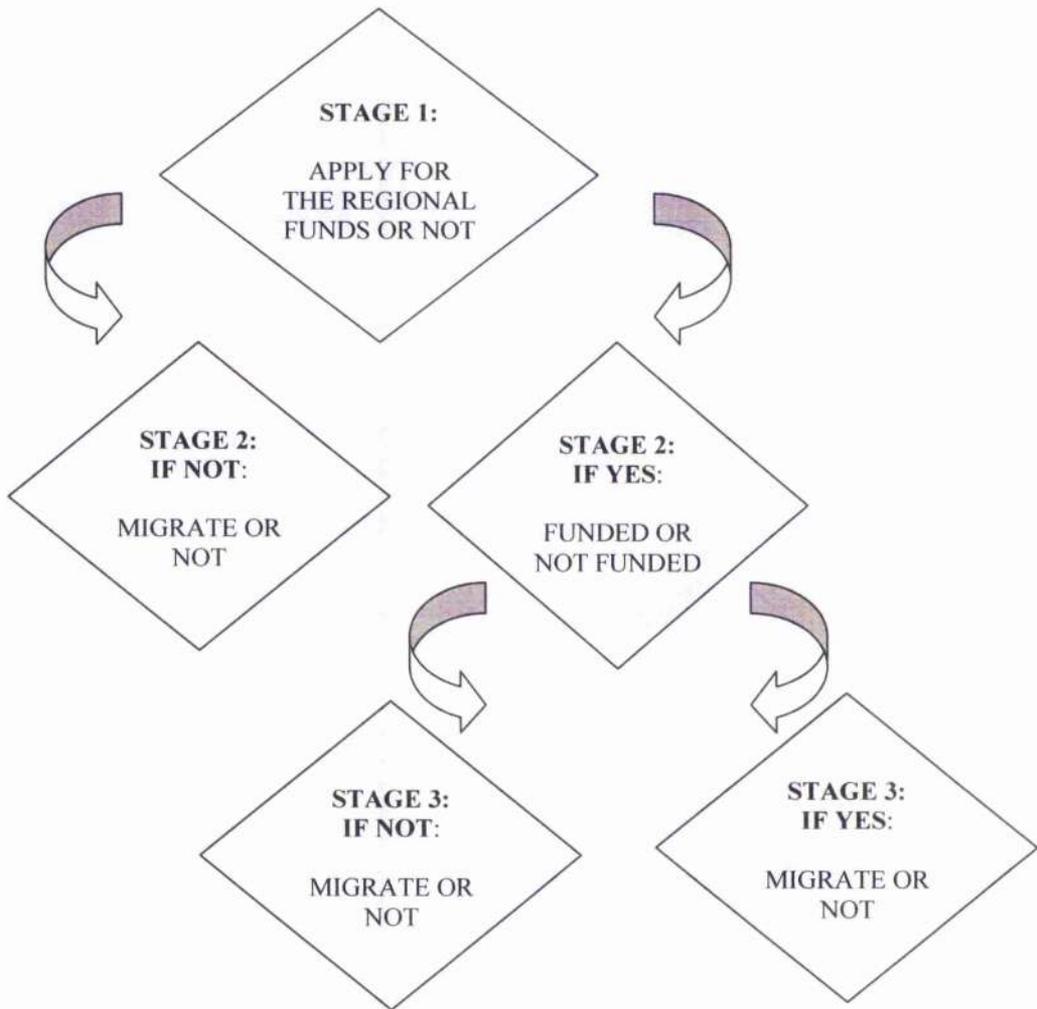
References:

Fink A. (1995), The survey handbook, in A. Fink (eds.), *The Survey Kit*

Foddy W. (1993), Constructing questions for interviews and questionnaires: theory and practice in social research, Cambridge: Cambridge University Press

Greenwood M. (1997), Internal Migration in Developed Countries, in M. R. Rosenzweig and O. Stark (eds.), *Handbook of Population and Family Economics* (Amsterdam: Elsevier), 1997, pp. 647-720

Lehtonen R. and Pahkinen E. J. (1995), Practical methods for design and analysis of complex surveys, Chichester: Wiley

Appendix 1 – Conduct of the survey: Cover Letter

Dear Name Surname,

We write to you with reference to the 'Borsa di Formazione', sponsored by Regione Basilicata, you benefited from in 1995 for enrolling a master course.

As you know, through 'Borse di Formazione', Regione Basilicata offers, to young and talented individuals of our region, financial support aimed at acquiring valuable expertise and skills.

Given the relevance of this policy measure and the considerable amount of financial resource employed, we believe that it is necessary to proceed with an accurate and comprehensive evaluation of its effectiveness and efficiency.

Important lessons can be learned directly from the beneficiaries of the policy measure. In particular, by paying the utmost attention to the observations and opinions about their experiences and by monitoring their subsequent job market performances.

For this purpose, our bureau has designed a questionnaire, which you will find herewith attached. The aim of the questionnaire is that of collecting information on the course you attended and your subsequent career history. All the information will be elaborated in anonymous form. We kindly ask you to complete the questionnaire and promptly return it, in one of the ways described below.

We would like to stress that the information we ask you to kindly provide, together with those of other beneficiaries, will enable us to deliver policy measures which improve more effectively the well-being of our local community.

Thank you for your effort and assistance.

Sincerely,

Doctor Vinicio Capossela
Director of Ufficio Formazione Professionale
Regione Basilicata

Important: in case the person concerned has changed place of residence, we should be most grateful if you would communicate the new address/telephone number or email address to our bureau. Please use your preferred following option to communicate with us:
- Mail (using the pre-stamped envelope)
- Fax n. +390000001 (to the attention of Ms. Bellini);
- Telephone n. +300000002 (Ms. Bellini)
- Email: Bellini@Regione.Basilicata.it

REGIONE BASILICATA

Appendix 2 - SURVEY QUESTIONNAIRE

The information you supply will be elaborated in anonymous form. Information will be treated in accordance with Legge n. 675 del 31/12/1996 "Tutela delle persone e di altri soggetti rispetto al trattamento dei dati personali".[**Privacy law**]

How to return the questionnaire:

- by mail: address to Regione Basilicata (using the freepost envelope provided)
- by fax: Regione Basilicata fax n. 0039/1111111

For help and further information please contact: Dott.ssa Stefania Bellini – Regione Basilicata – Dip.to Formazione – Corso Italia 1 – 1000 Potenza (PZ)

Tel. +39000001

Fax. +39000002

Email address: Bellini@RegioneBasilicata.it

Section (A) - Personal details

Name and Surname_____

Date of birth_____place of birth_____

Current address_____

_____

University degree obtained_____

year_____final mark_____at the University of_____ _____

_____

Are you married? ☐ YES, ☐ NO; (if **YES**, is your partner employed? ☐ YES ☐ NO)

Section (B) – Information on the course financed by Regione Basilicata

(1) Postgraduate course attended_____

_____

Educational Institution_____

Length (in months)_____ Location_____

(2) At that time, would you have enrolled in the course regardless the financial contribution of Regione Basilicata? ☐ YES ☐ NO

(3) What is your opinion on the overall quality of the course attended?
(*5 = very good, 4 = good, 3 = adequate, 2 = poor, 1 = very poor*)
☐ 5, ☐ 4, ☐ 3, ☐ 2, ☐ 1

(4) How do you think the course contributed to your subsequent placement in the job market?
(*3 = very important, 2 = useful, 1 = not relevant*)
  ☐ 3, ☐ 2, ☐ 1

(5) Have you obtained an internship as part of the master/course attended?
☐ YES ☐ NO

if **YES,** specify: Institution_____

Location_____Length in months_____

(6) What is your opinion on the experiences acquired during the internship period?
(*5 = very good, 4 = good, 3 = adequate, 2 = poor, 1 = very poor*)
☐ 5, ☐ 4, ☐ 3, ☐ 2, ☐ 1

Further remarks on the master/course attended_____

_____

Section (C) – Job market search and first employment after the course

(1) At the end of your course, which active steps have you taken to find a job?

You have......... (insert an *X* where appropriate)

_____ used the support of the same Institution where you attended the master/course

_____ used Internet

_____ replied to an advertisement

_____ asked friends or other contacts

_____ used the support of a Job Centre/Employment Agency or other public/private institution (describe_____)

_____taken steps to start your own business (describe_____)

_____other (describe_____)

(2) At the end of your course, which would have been your preferred location?

(1 = *first choice*; 2 = *second choice*)

_____ Basilicata

_____ Region in which you attended university studies

_____ Region in which you attended the financed course/master

_____ Another Italian region (specify_____)

_____ Abroad (specify_____)

_____ No preferences

(3) After how many months from completion of the course did you found your first employment?

☐ _____ (indicate the number of months)   ☐ I have never worked[1]

---

[1] **In this case the questionnaire for you ends here. Thank you for your time and effort.**

(4) At completion of your course, have you received more than one job proposal?

☐ YES   ☐ NO

if **YES,** indicate in which region:

in Basilicata ☐   in an other region ☐ (specify_____)

give a brief description of the job proposals_____

_____

(5) At the end of your course, which was your first employment?

(insert an **X** where appropriate)

(a) ☐ Self-employed

describe_____

Location_____

(b) ☐ Employee (permanent position) or ☐ Employee (short-term job)

at_____

Location_____ position covered_____

☐ Private company   ☐ Public sector

Date of start: month_____year_____

(7) Which of the following factors do you acknowledge to have been relevant in influencing your decision between alternative opportunities?
(insert an **X** where appropriate)

____ higher remuneration

____ better career prospects and opportunities

____ preferred geographical location

____ typology of job preferred

____ geographical location suitable for your partner and/or other memberst of your family

____ other reasons (specify_____)

Section (D) - Current employment position

(1) Describe your current employment position only if it is different from the one you described in the previous section

(insert an **X** where appropriate)

(a) ☐ Unemployed (since: month_____year_____)

(b) ☐ Self-employed describe_____

Location_____ type of work_____

(c) ☐ Employee (permanent position) or ☐ Employee (short-term job)

at_____

Location_____ position covered_____

☐ Private company   ☐ Public sector

Date of start: month_____year_____

(2) In which of the following income brackets is your yearly gross income included?

☐ less than € 20,000;   ☐ between € 20,000 and € 40,000; ☐

between € 40,000 and € 60,000;   ☐ more than € 60,000

(3) For how long had you been unemployed before starting the job described in this section?

_____ (indicate the number of months)

(4) Which of the following factors do you acknowledge to have been relevant in influencing the change from your previous job?

(insert an *X* where appropriate)

___ Employer's decision (layoff, end of short-term job contract, etc.)

___ Higher remuneration

___ Better career prospects and opportunities

___ Preferred geographical location

___ Typology of job preferred

___ Geographical location suitable for your partner or other members of your family

___ Other reasons (specify_____)

(Only if you currently do not work in Basilicata)

(5) In your opinion, which difficulties would you probably expect if you wanted to carry out your current type of work in our region?

(3 = *very relevant difficulty*; 2 = *fairly relevant*; 1 = *not relevant*)

☐ a lack of employment opportunities comparable to your current employment position;

☐ lower remuneration;

☐ lower quality of life for you and your family;

☐ geographical location not suitable for your partner or other members of your family

☐ other (specify_____)

(6) If you had the chance (i.e. obtaining a job position as preferred as your current one) would you be willing to move in Basilicata?

☐ YES   ☐ NO

(7) if **YES**, would you be willing to move, even accepting a lower remuneration?

☐ YES   ☐ NO

if **YES,** a reduction no lower than

☐ 5%,   ☐ 10%,   ☐15%,   ☐ >15%

Further considerations, comments or suggestions_____

_____

_____

_____

Important: if your first employment after the course (*Section C*) also represents your actual job status (*Section D*) and no **changes of job and/or location** have occurred, the questionnaire for you ends here. Otherwise continue with *Section (E)*.

Section (E) - Chronology of further relevant changes in your employment status

Note: in this section of the questionnaire you are kindly requested to give some information related to changes in job status and/or job location intercurred between your first and last employment.

Please describe these changes starting from the firts change after your first job (as reported in *Section (C)*) **until the last change before your current employment.**

Change n. 1

Have you changed your employment location?

☐ YES ☐ NO

(if **YES**, was the change of location due to your own decision? ☐ YES ☐ NO)

Have you started a new job?

☐ YES ☐ NO

if **YES**, specify the following information relative to the new job:

tipology_____

location_____

starting date:month_____year_____

For how long have you been unemployed before starting the new job?

_____ (indicate the number of months)

Which of the following factors do you acknowledge to have been relevant in influencing the change from your previous position?

(insert an *X* where appropriate)

___ Employer's decision (layoff, end of short-term job contract, etc.)

___ Higher remuneration

___ Better career prospects and opportunities

___ Preferred geographical location

___ Typology of job preferred

___ Geographical location suitable for your partner or other members of your family

___ Other reasons (please specify_____)

Change n. 2

Have you changed your employment location?

☐ YES   ☐ NO

(if **YES**, was the change of location due to your own decision? ☐ YES   ☐ NO)

Have you started a new job?

☐ YES   ☐ NO

if **YES**, specify the following information relative to the new job:

tipology_____

location_____

starting date:month_____year_____

For how long have you been unemployed before starting the new job?

_____ (indicate the number of months)

Which of the following factors do you acknowledge to have been relevant in influencing the change from your previous position?

(insert an *X* where appropriate)

___ Employer's decision (layoff, end of short-term job contract, etc.)

___ Higher remuneration

___ Better career prospects and opportunities

___ Preferred geographical location

___ Typology of job preferred

___ Geographical location suitable for your partner or other members of your family

___ Other reasons (please specify_____)

Other changes

_____

_____

_____

_____

_____

THANK YOU FOR YOUR TIME AND EFFORT

IN CASE YOU ARE INTERESTED IN RECEIVING THE RESULTS OF THIS SURVEY, WE WILL BE PLEASED TO SEND THE FINAL REPORT TO YOUR E_MAIL ADDRESS

_____

Chapter 5

Migropolis: networks and formation of ethnic communities in cities

Nicola D. Coniglio[*]

1. Introduction

Several theoretical contributions have attempted to explain the geographical concentration of migrants from the same source region which is a frequently observed pattern of migration experiences. While a large set of potential destinations exist, migration is often only channelled into a very limited number where migrants networks are established already. Carrington et al. (1996) and Chau (1997) both provide a dynamic model of labour migration in which moving costs decrease with the number of established migrants in a particular destination. They show that once started, migration flows from the same source area accelerate even while inter-regional income gaps are reduced. In addition, in many empirical studies of migration, the stock of established migrants from the same area/country is almost always found to be a positive and significant explanatory variable. Moving costs are endogenous and decreasing in the size of past migration flows and subsequent waves of migration benefit from information and support provided by established migrants.[1]

While the passive aspect of this migration dynamic has been previously recognised in the literature, the active role of established migrants has been largely ignored. Why do established migrants give support to newcomers? Given the considerable amount of effort often involved in these support activities, and given that support is not generally confined to family members but is often a community wide phenomenon, we have reason to believe that pure altruism is an important explanatory variable, but not the only one.[2]

---

[1] For empirical studies on migration networks which use micro-data level see Winters et al (2001), Munshi (2003), Chiswick and Miller (2002), Bauer et al (2002).

[2] Some recent contributions have emphasised the importance not only of *"strong ties"* such as households components or kin but also of *"weak ties"* as acquaintances and co-villagers. See Wilson (1998) on the importance of weak ties in Mexican migration networks in the US and Grannovetter (1973, 1982) on the importance of weak ties, in general, in locating employment opportunities. In their empirical study on Mexican migration networks in the US, Winters et al (2001) are able to disentangle the relative effects of strong (family) and weak (community) ties on migration propensity and migrants' location decision.

In this paper we extend previous models of chain migration by considering migration as a community-wide phenomenon where established migrants strategically provide support to newcomers. This active role of support provision is associated with positive externalities which new incoming migrants might have on migrants already settled in the host country.

To the best of our knowledge, the work by Stark and Wang (2002) is the only one to deal with the same issue. In their model, support to newcomers in a particular location is provided only by highly skilled established migrants motivated not by altruistic consideration but by pure self-interest. High-skilled migrants have an incentive to attract low-skilled migrants since working with a larger number of these individuals renders their skills relatively more scarce (and therefore more valuable), which in turn has a positive effect on their wages. Their approach postulates the existence of a "migrants' production function" where high and low-skilled migrants are jointly employed. Once in the host location, low-skilled migrants become high skilled in the next period through an on-the-job skill enhancement process. In turn, they will provide support to a new wave of low skilled migrants giving rise to a migration chain. In our opinion, this approach has some shortcomings. Firstly, it is not obvious why established low-skilled migrants would not use remittances strategically in order to keep potential new-comers at home (as in Stark 1999 or Docquier and Rapoport 1998). In addition, migration chains are a phenomena which take place within components of the same source location community. A model which tries to explain migration chains on the basis of production externalities, as Stark and Wang 2002, would not *a priori* exclude the possibility to observe natives or high-skilled migrants giving support to low-skilled migrants belonging to different ethnic groups, such as Indian and Chinese high-skilled workers giving support to Mexican low-skilled workers. Such a proposition is not empirically supported in migration studies. Furthermore, support by established migrants is not confined to high-skilled migrants providing help to low-skilled

ones. Support appears to go beyond differences in skills. Support is often more pronounced between low-skilled migrants directly competing in the host locality labour market, as was the case for Italians and Irish immigrants in the US at the beginning of the last century. In a recent study on the contribution of Chinese and Indians scientists and engineers to the Silicon Valley economy, Saxenian (1999) documents the existence of dense ethnic networks within these highly skilled immigrants.[3]

In this paper, we view migration networks not merely as an aseptic instrument to facilitate immediate settlement of individual migrants but as the manifestation of a long-term development of a new community (or reconstruction of the lost one). A quote from a study on migration chains of Italians in Australia conducted by Lever-Tracy and Holton (2001) is emblematic of the main idea of this paper:

"...He (*an established migrant who helped many newcomers*) just wanted as many Italians as possible in Australia...The more Italians came out the less foreign one would feel." (page 94)

We propose a model that explains both established migrants' support to newcomers and the formation of ethnic communities on the basis of the existence of "ethnic consumption externalities". Incoming migrants from the same ethnic community have an expansionistic effect on the number of ethnic-specific non-tradable goods and services available to established migrants in the host region. Culturally-based tastes for particular goods and services generate an ethnic consumer demand and only individuals from the same ethnic community have the skills or the "insider's information" required to provide these goods (protected market). If the ethnic population is large enough, an ethnic sector will emerge and

---

[3] Ethnic networks for these professionals are often institutionalised into professional associations. Virtually no overlap exists between different ethnic groups. These organisations combine elements of traditional immigrant culture and professional networking and exchange of information. They play a key role in supporting new immigrants in their professional and social adjustment in the US.

eventually grow as the population expands further. Migrant communities in almost every big city around the world provide a wide range of ethnic goods and services. Ethnic specific goods and services include for instance preparation of ethnic foods, banking and financial services, ethnic media such as magazines, radio, newspapers, religious services, organised unions, political groups and usually a vast array of migrant associations.

As the size of the ethnic community grows larger and larger, discrimination in the labour market, prejudice and hostility from the larger society generate negative externalities which might offset any positive effect. For instance, during the Great Migration, African Americans in northern US cities became increasingly visible and as a consequence racist acts and violence against them increased. Congestion costs as the community expands might also be linked to the existence of a limited stock of housing or to increased competition in the labour market, compressing wages.

The trade-off between these two opposing forces implies the existence of an optimal size of an ethnic community. In a recent contribution, Bauer et al. (2002) find evidence of an inverse U-shaped relationship between the number of Mexican migrants from the same village of origin to a particular US location and the probability that subsequent migrants from the same village will also choose that destination. Too many immigrants from the same village in a particular location decreases the probability of choosing that location.

Although migration chains and the formation of ethnic communities characterise immigration and emigration experiences of many countries (both developed and LDCs) across all periods of human history, not all migration flows follow this pattern. Even more interesting is the fact that in this regard ethnic groups sometimes have different migratory behaviour.[4] In our model, according to the degree of preferences toward ethnic consumption,

---

[4] Ethnic concentration of migrants is not a phenomenon characterising only migration flows of low-skilled migrants from poor countries, as it is sometimes believed. For instance, a study by Glebe (1986) shows that the Japanese community in Dusseldorf is the most concentrated of all minorities in that city. White (1998) conducts a related study on the pattern of settlement of developed country migrants' in London. Jewish communities

the mobility costs of the source country population and congestion costs (hostility externalities) in the host location, different scenarios arise. These scenarios provide a possible explanation for variations in the observed migration patterns of different ethnic communities.[5]

In section 2 we provide a brief review of recent empirical work on migration networks. A discussion on the definition and nature of ethnic goods is presented in section 3. In section 4 we outline the basic structure of the model and analyse the interaction between the preference for ethnic consumption and the size of the ethnic community in a given host location. We derive the conditions under which strategic support by established migrants is observed and the implications on the size of the ethnic community. In the same section alternative scenarios are outlined. Section 5 concludes.

2. Migration networks, the role of established migrants and new immigrants' location decision

Established migrants often provide direct assistance to new migrants. They might help finance transportation costs, provide housing or pre-arrange jobs. In addition, established migrants in the network supply information, which allows recipients to assess more carefully the expected returns from migration. Nelson (1959) has classified information on which potential migrants form their expectations into two main categories: specific and generalised information. Specific information implies knowledge about particular opportunities, while generalised information concern the awareness about the attractiveness of opportunities in a location. Established migrants have a fundamental role in conveying both types of

---

around the world represent another example of ethnic concentration mainly aimed at preserving the community identity and an ethnic social capital.
[5] For empirical evidence on the clustering patterns of immigrants by country of origin or ethnic group see Zavodny (1999) and Chiswick et al (2002) for the US and Australia respectively.

information. In particular, by providing specific information and direct assistance they considerably reduce migration costs for incoming migrants.

In his study of rural-to-urban migration in India, Banerjee (1984) finds that over one-half of the migrants left their place of origin because of suggestions to migrate received from contacts (relatives of friends) settled in the urban area. A considerable share of these migrants had a pre-arranged job or received assurances of urban employment before leaving the rural area. Menjivar (1995) shows that for newly arrived Mexican migrants in California it is common to be hosted by kin (relatives but also members of their hometown), borrowing money and receive other important form of assistance from them.

The availability of migration networks in certain host locations significantly affects individuals' migration decisions. The probability of migration to the US is higher for Mexican households with prior migration experience and for households originating from communities with extensive migration history (Massey and Garcia Espana, 1987).[6]

Migration networks, and the resulting exchange of information and support, are not only the result of **"strong ties"** such as those within households or kin but also of **"weak ties"** such as acquaintances and co-villagers.[7] In a recent article, Winters et al. (2001) find strong support of the positive influence of migrants networks on (i) the decision to migrate; (ii) where to migrate and (iii) the number of migrants sent by a Mexican household. In their investigation, they are able to disentangle the separate effects on migration probability of strong (family) and weak (community) ties. They find that family and community networks are substitutes in the generation of information and general assistance. Households with weaker family ties derive more benefit from community networks than households with larger family networks.

---

[6] On the consequences of migration networks in terms of labour market performance and immigrants adjustements see Munshi (2003), Chiswick and Miller (2002), Gonzales (1998), Borjas (1995).

[7] See Wilson (1998) on the importance of weak ties in Mexican migration networks in the US and Grannovetter (1973, 1982) on the importance of weak ties, in general, in locating employment opportunities.

Taylor (1986), in a model of household labour allocations under uncertainty emphasises the risk-reducing role of migration networks. Information, transmitted by established migrants, can improve the allocative efficiency of households' labour decisions. In fact, their perception of the distribution of returns (*subjective distribution*) will match more accurately the true distribution (*objective distribution*). The value of accurate information is higher in the context of international migration, which is characterised by a high-risk environment, larger sunk costs (transportation, adjustment costs, psychical costs etc.) and scarcity of information. Migration networks are expected to have a greater role in international than in internal migration. Econometric analysis using data collected in two Mexican villages confirms the predictions of Taylor's model.

An additional explanation to immigrant clustering is given by the existence of "herd behaviour" in migration. Epstein (2002) shows how potential migrants, who have some private information but are imperfectly informed about the attributes of alternative locations, might rationally decide to discount their private information and replicate the decisions taken by previous migrants. In his model, individuals are able to observe previous emigrants' decisions but not the information signal that was driving those decisions. Herd behaviour might result in an inefficient equilibrium. Migrants may regret the choice taken and would have gained higher utility had they not followed the herd and chosen a different destination. This theory, based on "information cascades" between subsequent waves of migrants, should not be considered as a completely different explanation for clustering. Herd behaviour and networks externalities might clearly coexist. In our opinion, information cascades in Epstein's theory are equivalent to the notion of "generic information" flows as developed by Nelson (1959). Generic information is often widespread without the active role or even the will of previous migrants. For instance, information about the locations chosen by people moving out of our village might be public knowledge. We might learn about it by talking

with the butcher or barber of the village or sipping a coffee in the main square's café. We can even form our suppositions on how successful a migrant is by looking, for instance, at his/her remittances home. Even if not directly observable, migrant's riches can be assessed by observing the consumption of the remittances' beneficiaries. This public knowledge about migrants' economic performances in a given destination might certainly induce "herd behaviour" and therefore subsequent migration waves, which are not triggered by an active role of established migrants.

Our model is based on a different premise. The information cascade theory explains immigrants clustering on the basis of the existence of imperfect information in a framework where established migrants have no active role in "recruiting" newcomers. In our theory, established migrants might play a very important role and their strategic decision whether or not to provide support to newcomers is linked to the interplay between positive ethnic consumption externalities and negative congestion costs. We are able to depict a richer framework which explains not only the formation of ethnic concentration but also the development of an ethnic productive sector in a particular location. Production of a rich set of ethnic goods and services is a striking feature of the various Chinatowns, Little Italys or Little Karachis around the world.

3. On ethnic goods and services

Consumption preference is one of the dimensions in which migrants often differ from the native population in the host locality. Immigrants from a particular ethnic group might have distinctive preferences toward a set of consumption varieties (which we define "ethnic goods and services"), not shared with the host population. Distinctiveness in consumption is particularly strong for ethnic migrants coming from source communities whose individuals

132

share a strong cultural identity. Ethnic varieties are broadly defined to include both market and non-market goods and services. In the latter category will fall, for instance, social interactions between member of the same community of origin as normal gatherings and celebration of certain specific religious festivals. The ethnic identity of these goods is linked to the fact that they address needs specific to the individual belonging to the ethnic community. In the consumption bundle of immigrants, ethnic goods and services might be of considerable importance. As a consequence, availability and prices of these goods in the host location might play an important role in shaping the location decision of ethnic migrants.

In order to understand how ethnic consumption might affect immigrants' location decisions and the migration of an ethnic community, it is important to define first, the nature of ethnic commodities.

Some of these goods are **tradable**. Ethnic clothing, particular ingredients essential for preparing ethnic meals, CDs of Italian, Norwegian or Pakistani singers and DVDs of Tibetan or Chinese movies might all be traded over long distances. Availability of these tradable goods is, in principle, identical across potential localities in the host destination. Prices at different locations are however unlikely to be identical. Firstly, transportation costs might differ. Although for a single commodity transport costs might not be highly significant, the sum of these costs across the entire bundle of ethnic consumption might be considerable. If potential destinations are in different countries, duties and import taxes levied by national authorities, might contribute to price disparities. Secondly, and probably more significant, a certain amount of fixed costs are associated with the distribution of ethnic goods. Setting up a distribution channel is costly. As a consequence, the equilibrium price of an ethnic variety will be higher in locations with a smaller number of ethnic consumers. In addition, localities with a small number of ethnic consumers are likely to have only one or a limited number of suppliers/retailers of a particular product. The lack of competition in the retailing sector is

likely to translate into higher prices. For example, if parmesan cheese is sold only in one shop in town, the shopkeeper will be able to extract greater surplus by acting as a local monopoly than if competition was fierce. In summary, for all these reasons, the price of a given bundle of consumption of tradable ethnic commodities might be decreasing in the size of the ethnic community in the host location.

A considerable share of ethnic goods and services are **non-tradable**. Demand for a type of religious service or the service of a doctor applying traditional Chinese medicine both fall into this category. When the degree of social and cultural distance between the ethnic group and the host population is high, goods and services which are not *per se* ethnic might be considered so, in virtue of the fact that they are provided by members of the same ethnic group of the consumer. For example the service offered by a lawyer appear in general to have no distinctive ethnic component. Nevertheless, a Mexican immigrant in Los Angeles may prefer to use the services of a Hispanic lawyer than those of a non-Hispanic lawyer. The Mexican customer might have special consumer demands that a lawyer with an insider knowledge of the ethnic's group culture may be able to provide more effectively. Consequently, even if the price charged and the final results of the service are identical, the Mexican customer will probably find himself more at ease with a member of his own community and hence more likely to use their services. Non-pecuniary transaction costs in the trade between the lawyer and his client, such as language barriers, might be considerably reduced when these individuals belong to the same community of origin. The number of ethnic non-tradable varieties available in a particular host location is largely dependent on the intensity of preferences for ethnic consumption and on the size of the ethnic group. If the group is sufficiently large, it may support the existence in the host location of a differentiated

set of ethnic varieties. In many large cities across the world, immigration transforms some neighbourhoods into replicas of towns or communities of the source country.[8]

If ethnic consumption has a high value, a migrant will prefer, ceteris paribus, a location which offers a wide selection of ethnic goods at a lower price, i.e. a location with a large concentration of migrants from the same ethnic community.

Given the intrinsic nature of ethnic commodities, individuals belonging to the same ethnic group as the consumers are best suited at producing them. In practice, insider's knowledge of the ethnic community's culture is essential in providing goods and services closer to consumer tastes. Ethnic producers tend to have to a certain extent a protected market, i.e. a competitive edge over non-ethnic producers. In some cases, for certain goods and services, the market can be completely closed and ethnic producers are not only more able than others to satisfy demand but are in fact the only ones who can actually supply the ethnic market. An example, illustrated by Boyd (1998), is that of religious services demanded by African-American immigrants from the American South in urban northern areas during the Great Black Migration in the US. What was "ethnic" about these religious services? African Americans "were accustomed to services accompanied by improvisational singing, shouting and other form of active participation and demonstrative enthusiasm". Pioneer migrants were dissatisfied with "intellectual sermons" and cold and impersonal large congregations they found in the northern cities (Grossman, 1989). Large flows of southern migrants created the demand for churches and therefore, pastors travelling northward, who were better able to accommodate their special tastes.

From the above discussion it is clear that the number of both tradable and non-tradable varieties is highly likely to be positively affected by the size of the ethnic group in the host

---

[8] For an interesting sociological study on Hungarian immigrants in North America and their pattern of location/ethnic consumption during the first half of 20th century see Kosa (1956).

location. In what follows, we capture this idea in the simplest possible way by considering the existence of non-tradable varieties only.

4. The model

In order to analyse the relationship between individuals' migration decisions, migration chains and preferences for non-tradable ethnic goods and services, we specify a simple model representing a single potential destination. This section is divided into three parts. In the first part the migration behaviour and consumption preferences are presented. The production structure of the economy is specified in the second part, while in the last part, migration pattern and the optimal size of the ethnic group are analysed.

4.1 Migrant's utility

A migrant's utility is defined over two types of goods and services, ethnic (E) and native (Y). In line with our discussion above, we assume that, production of ethnic goods requires the employment of individuals of the same ethnic group as the consumers (i.e. the same source location). We assume that the goods and services are non-tradable, this means that the number of ethnic varieties available depends on the size of the ethnic group in a given locality.

Native goods are purely labour intensive (i.e. labour is the only input), where the labour pool consists of both migrants and natives. The two types of goods in the economy are produced in many differentiated varieties. We define n_E and n_Y as the number of respectively type E and type Y goods.

Utility of a representative migrant can be expressed by the following function:[9]

$$V = \int_{i=0}^{n_E} ((1+\varepsilon)c_{Ei})^\theta \, di + \int_{j=0}^{n_Y} (c_{Yj})^\theta \, dj - \xi M \qquad (1)$$

where $\varepsilon > 0$ indicates that ethnic migrants particularly enjoy consumption of ethnic varieties. This formulation of the utility function implies that ethnic and non-ethnic goods and services are substitutes. Ethnic producers have "insider information" and particular skills which enable them to provide products closer to the taste and needs of ethnic consumers. As highlighted above, this might include the cooking of a particular meal, provision of religious services or simply the ordering and display of goods in the way it is done in the source country. The parameter $\theta \in (0,1)$ captures the intensity for the love of variety in the utility function; as $\theta \to 0$, consumers derive more utility from product variety and as $\theta \to 1$, they derive less utility from product variety. c_{Ei} and c_{Yj} represent respectively consumption of good i, j of type E and Y. The last term on the right-hand-side (ξM) captures the fact that as the size of the ethnic group in one location increases (M), i.e. as the ethnic community becomes more "visible", migrants suffer from hostility and/or discrimination inflicted upon them by native inhabitants. For simplicity we assume a linear "hostility externality" measured by $\xi \geq 0$ times the size of the ethnic group M.[10] More generally, the hostility externality captures congestion costs of a different nature associated with a dimensional expansion of the ethnic community.

---

[9] In what follows we use a modelling strategy for the migrants' preferences similar to that employed by Glazer et al. (2003) in a different context.

[10] For an interesting theoretical discussion concerning the "hostility externality" suffered by minority groups see Carlton (1995). Another interesting reading on hostility and the size of a minority is Rapoport and Weiss (2003).

We assume that each migrant is endowed with one unit of labour which is inelastically supplied in the individual's location of residence. The budget constraint of a representative migrant is given by:

$$w = \int_{i=0}^{n_E} p_E(i) c_E(i) \, di + \int_{j=0}^{n_Y} p_Y(j) c_Y(j) \, dj \qquad (2)$$

From equations (1) and (2) we obtain the following migrant's demand for a single E and Y variety (*see Appendix I*):

$$c_E = \frac{w}{n_E p_E + n_Y p_Y \varepsilon'};$$
$$c_Y = \frac{w \varepsilon'}{n_E p_E + n_Y p_Y \varepsilon'} \qquad (3)$$

where $\varepsilon' = (1/(1+\varepsilon))^{\theta/(1-\theta)}$. Conversely, native individuals' demand for a single native variety might be expressed as:

$$c_Y^{nat} = \frac{w}{n_Y p_Y} \qquad (4)$$

where superscript '*nat*' identifies native variables. Native individuals do not benefit from ethnic varieties.

4.2 Production technology

The two sectors of our economy have the same production technology. The only difference consists of the fact that the ethnic sector employs a specific factor of production, 'ethnic labour'. Since all varieties, both of type E and type Y, are symmetric in what follows we consider production of a representative firm irrespective of the type of products. Each variety is produced at decreasing average costs. An initial effort (fixed costs) is needed to set up a new business, and then each unit is produced at constant marginal cost. The labour input requirement for each variety is given by:

$$l_i = \alpha + \beta x_i \tag{5}$$

where α, β are respectively the fixed and the variable costs and x_i is the output level of the differentiated good or service. Love of variety and increasing returns at the firm level imply that each newcomer in the non-tradable sector will produce a new variety in order to capture some 'monopolistic power'.

The profit-maximising price is a mark-up over the marginal costs (wage):

$$p_i = \frac{\beta w}{\theta} \tag{6}$$

As long as excess profits exist, firms will continue to enter the market until profits are driven to zero, that is $\pi_i = p_i x_i - (\alpha + \beta x_i) w = 0$. It follows that in equilibrium the optimal quantity produced by each firm is:

$$x_i^* = \frac{\alpha}{\beta}\left(\frac{\theta}{1-\theta}\right) \tag{7}$$

Substituting (7) into eq. (5) we obtain the optimal labour input requirement for each firm:

$$l_i^* = \alpha + \beta x_i^* = \alpha + \beta \frac{\alpha}{\beta}\left(\frac{\theta}{1-\theta}\right) = \frac{\alpha}{1-\theta} \tag{8}$$

4.3 Size of the ethnic group and number of ethnic varieties

Migrants can be employed in either Y or E sectors. In addition, they can move freely between sectors, equalising wages in equilibrium. The number of ethnic and native varieties produced in equilibrium is demand driven (given by the share of income spent on ethnic and native goods and services).

When will an ethnic sector emerge? The first ethnic variety will be produced in the host location only if there is sufficient demand for it. In other words, migrants demand (Mc_E) is sufficiently large to cover the break-even optimal quantity x_E^* as defined by eq. (7). If this condition is not satisfied, i.e. the size of the migrants community is below a critical mass, migrants will be employed in the production of native varieties and consume only native products. In this case, migrants will have the following utility and their individual demand will be exactly as that of natives as defined by eq. (4):

$$V^{no\_ethnic} = \int_{j=0}^{n_y} c_Y(j)^\theta \, dj - \xi M \tag{9}$$

Equilibrium implies that goods and labour markets clear, that is, demand equals supply for each variety and all individuals are employed. When only the production of native varieties takes place equilibrium in the goods market is given by the following expression:

$$M\frac{w}{n_Y p_Y} + L\frac{w}{n_Y p_Y} = x_Y^* \Rightarrow \frac{(M+L)w}{n_Y p_Y} = \frac{\alpha}{\beta}\left(\frac{\theta}{1-\theta}\right) \tag{10}$$

where M, L represent the size of the ethnic and native populations. By replacing eq. (6) into (10) we find the equilibrium number of native varieties when an ethnic productive sector does not exist:

$$n_Y^{no-ethnic} = (M+L)\left(\frac{1-\theta}{\alpha}\right) \tag{11}$$

In a similar fashion, when the size of the ethnic community is large enough to sustain an ethnic productive sector, the equilibrium conditions in the goods markets for a representative ethnic (E) and native (Y) variety are as follows:

$$Mc_E = x_E \Rightarrow \frac{Mw}{n_E p_E + n_Y p_Y \varepsilon'} = \frac{\alpha}{\beta}\left(\frac{\theta}{1-\theta}\right) \tag{12}$$

$$Mc_Y + L\frac{w}{n_Y p_Y} = x_Y \Rightarrow \frac{Mw\varepsilon'}{n_E p_E + n_Y p_Y \varepsilon'} + \frac{Lw}{n_Y p_Y} = \frac{\alpha}{\beta}\left(\frac{\theta}{1-\theta}\right)$$

where we assume that natives' consumption of ethnic variety is equal to zero.

Substituting the equilibrium prices $p_E = p_Y = \beta w / \theta$ in (12) we obtain the following system:

$$\begin{cases} n_E = \dfrac{M\left(1-\theta\right)}{\alpha} - n_Y \varepsilon' \\[2mm] \dfrac{M\varepsilon'}{n_E + n_Y \varepsilon'} + \dfrac{L}{n_Y} = \dfrac{\alpha}{1-\theta} \end{cases}$$

Solving these two equations yields the equilibrium number of varieties:

$$n_E = \left(\frac{1-\theta}{\alpha}\right)\left[M - \frac{L\varepsilon'}{1-\varepsilon'}\right]$$

(13)

$$n_Y = \left(\frac{1-\theta}{\alpha}\right)\frac{L}{1-\varepsilon'}$$

An ethnic productive sector will emerge ($n_E > 0$) if and only if the following condition is satisfied:

Condition 1: *An ethnic sector exists if and only if the share of migrants from the same ethnic community with respect to the native population (M / L) is larger than $\varepsilon'/\left(1-\varepsilon'\right)$.*

We define $\widetilde{M}$ as this threshold size of the ethnic community:

$$\widetilde{M} = \frac{L\varepsilon'}{1-\varepsilon'}$$

(14)

For a given level of the native population (*L*), the stronger migrants' preference toward ethnic variety *vis-à-vis* native ones ($\varepsilon \rightarrow large$), the lower the critical mass of the migrants' community for which an ethnic sector arises. In addition, the lower the intensity of the love-for-variety ($\theta \rightarrow 1$), the lower will be $\widetilde{M}$. The intuition behind this last result follows from

142

the fact that the higher θ, the lower the gain in utility due to diversification of consumption. As a consequence, the ethnic migrant will shift expenditure toward consumption of the goods which provide him or her with the highest level of utility, i.e. ethnic varieties.

Proof: Ethnic and native variety use the same technology, i.e. the same labour input requirement l_i^a. Given the amount of total labour available in the host location economy $(M+L)$, the following relationship holds: $n_Y^{no-ethnic} = n_Y + n_E = (M+L)/l_i^*$. As long as migrants in the host location find it more profitable to demand only native varieties no ethnic goods will be produced ($V^{no-ethnic} > V$). Only when the marginal utility from having an ethnic variety is equal to that of having an extra native one, will ethnic production start. The critical mass of migrants necessary to sustain an ethnic productive sector is found by solving the following equality $V^{no-ethnic} = V$. Using eqs. (1), (9), (11) and (13) and solving for M, we obtain eq. (14). (see appendix II) ■

4.4 The pattern of migration and the optimal size of the ethnic group

4.4.1 The autonomous flow of migrants

All workers in the source location have the same level of skills and inelastically supply one unit of labour for which they are paid a wage rate (w_s), which is only a fraction of the wage offered in the host location (w_h). However, migration is costly. In the source location, each individual j in the population (P) faces a mobility cost a_j, which is broadly defined to include the full cost of relocating in the new destination such as psychological costs, travel expenditure and the considerable amount of information required in order to take the decision

143

to migrate. Individuals face different mobility costs, which we assume to be continuously distributed on the interval, $a_j \in \left[\underline{a}, \overline{a}\right]$, where $\underline{a}$ and $\overline{a}$ identify the most mobile and more immobile individual in the population respectively.[11] The fraction of population in the source region with a mobility cost equal to or lower than a_j is given by $F(a_j) = \int_{\underline{a}}^{a_j} f(a)Pda$, where $f(a)$ is the probability density function for any particular level of migration cost a. $F(a)$ is, therefore, the cumulative distribution function and is strictly positive, increasing and differentiable. A very simple and tractable case, which we will use in the following analysis, is when mobility costs are uniformly distributed, that is $f(a_j) = 1/\left[\overline{a} - \underline{a}\right]$ for any $a_j \in \left[\underline{a}, \overline{a}\right]$.

On the basis of available information on potential destinations, an individual will move toward the destination (if any) which maximises his/her utility net of migration costs. The first individual (or group of individuals) will migrate if utility in the host location net of migration costs is higher than utility at home. Formally, if $V_h - V_s > a_j$ individual j will migrate, where V_h and V_s are indirect utility levels in the host and source destinations respectively. Without loss of generality and only for analytical convenience, we assume that the availability of ethnic varieties in the source location is fixed at a given level and migration is relatively small so that it does not have any detrimental effect on the number of varieties produced at home. This implies that we can treat the utility associated with the source locality as fixed at a constant level $\overline{V}_s$.[12]

---

[11] These captures the fact that some individuals are more mobile than others for factors such as personal and household characteristics (age, gender, homeowners etc.), the degree of involvement in community life in the origin region and exogenously given preferences over the home location.

[12] We acknowledge that migration flows sometimes might have detrimental effects on the source economy. That is particularly true if migration is highly selective towards more skilled individuals, as is often the case. In our model we rule out this consideration since we are exclusively interested in investigating the pattern of migration flows rather than the consequences thereof on those left behind in the source location.

In our simple two-locations model, the initial autonomous flow of migrants' (M_1) depends on the utility differential (explained by both the exogenous wage rate differential and the endogenous availability/composition of non-tradable varieties) and the distribution among the population of 'mobility costs' in the source location. For a given set of expected host/source utility levels, the *marginal migrant* of the initial migration flow (M_1), which is indifferent between migrating or staying at home has the following mobility cost:[13]

$$a_1 = V_h(M_1) - \overline{V_s} \tag{15}$$

The corresponding size of the initial ethnic migrant community is given by:

$$M_1 = \int_{\underline{a}}^{a_1} f(a) P da \tag{16}$$

A corner solution with no migration (complete de-population) arises when equation (15) is never satisfied and $a_j > V_h(M) - \overline{V_s}$ ($a_j < V_h(M) - \overline{V_s}$) for any $a_j \in \left[\underline{a}, \overline{a}\right]$. In what follows we concentrate our attention on the pattern of migration in the case of an internal solution, $0 < M_1 < P$, i.e. a positive and finite initial flow of migrants.

Since ethnic goods are initially not available in the host destination, a relatively large wage gap could be necessary to kick-start migration. The most mobile individuals of the source location (if migration takes place) will compose the initial flow of migrants.

Equation (16) measures the size of the ethnic community in the host locality without strategic support from established migrants to potential incoming migrants.

---

[13] While the setting of our model is fundamentally static, its interpretation is implicitly dynamic as it is frequently the case in the economic literature aimed at analysing the spatial structure of the economy.

In what follows, we analyse the effect of incoming migrants from the same source community on the utility of established migrants. In addition, we show how the active role of established migrants in supporting newcomers enables us to depict a rich set of possible scenarios in terms of patterns of migration and ethnic community size in the host locality.

4.4.2 The effect of newcomers on established migrants' utility and the optimal size of the ethnic community

Once in the new location, established migrants have an incentive to give support to newcomers only if new migrants have a positive effect on their utility. Potential gains from incoming migration in terms of utility are endogenous and depend on the size of the resulting ethnic community in the host location. In the context of our model, different scenarios may emerge according to the size of the initial, autonomous flow of migrants (M_1), the degree of migrants' preferences toward the ethnic goods and services (ε), and the reaction of natives to increasing immigration flows (hostility externality, ξM).

Ethnic varieties will be available in the host localities only if *condition 1* is satisfied, that is a critical mass $\widetilde{M}$ of ethnic migrants is already settled in the host locality. The effects of incoming migrants will therefore differ according to the existence (or not) of an ethnic productive sector. Two cases emerge.

Case 1. **No ethnic productive sector** (size of the initial pool of migrants $\widetilde{M} \geq M_1$)

If the size of the initial ethnic community is smaller than the critical mass needed to establish an ethnic sector, migrants will initially consume only native non-tradable goods and services. Indirect utility in this case is:

$$V_h^{**} = (L+M)^{1-\theta} \left(\frac{\theta}{\beta}\right)^{\theta} \left(\frac{1-\theta}{\alpha}\right)^{1-\theta} - \xi M \qquad (17)$$

From the above equation it follows that the effect of further immigration on the utility of the established migrants depends on the trade-off between the positive *native variety expansion effect* and the negative effect due to increasing *ethnic discrimination/congestion costs*.

The optimal size of the migrant community is reached when the two opposite effects are balanced, i.e. $\partial V / \partial M = 0$. We call this threshold level M^{**}, which is given by the following expression:

$$M^{**} = \left(\frac{\theta}{\beta}\right)(1-\theta)^{\frac{2-\theta}{\theta}} \alpha^{-\frac{1-\theta}{\theta}} \xi^{-\frac{1}{\theta}} - L \qquad (18)$$

This size of the ethnic community is realised when only all individuals with mobility cost $a < a_j \leq a^{**}$ find it convenient to migrate, where a^{**} is the threshold level of mobility cost for which $M^{**} = \int_{\underline{a}}^{a^{**}} f(a) P da$

Case 2. **Ethnic productive sector** (size of the initial pool of migrants $M_1 > \widetilde{M}$)

When the size of the community is larger than or equal to $\widetilde{M}$, an ethnic productive sector is set up. Indirect utility of the established migrants (and for incoming migrants) in this case is:

$$V_h^* = M^{1-\theta}\left(\frac{\theta(1+\varepsilon)}{\beta}\right)^{\theta}\left(\frac{1-\theta}{\alpha}\right)^{1-\theta} - \xi M \tag{19}$$

Also in this situation a trade-off associated with subsequent immigration exists. The only difference is in the nature of the positive externality, which in this second case is due to an *ethnic variety expansion effect.*

Also for this case, in which the initial autonomous flow of migrants is above $\widetilde{M}$, we compute the optimal size of the ethnic community:

$$M^\circ = \left(\frac{\theta(1+\varepsilon)}{\beta}\right)(1-\theta)^{\frac{2-\theta}{\theta}}\alpha^{-\frac{1-\theta}{\theta}}\xi^{-\frac{1}{\theta}} \tag{20}$$

We define a° as the level of mobility cost associated with $M^\circ = \int_a^{a^\circ} f(a)Pda$.

The optimal size of the ethnic community in *Case 1* is always smaller than in *Case 2*, i.e. $M^* > M^{**}$ (see Appendix III).

Figure 1 depicts the level of utility that migrants enjoy in the host location (net of the constant level enjoyed in the source country $V_{h-s} = V_h - \overline{V_s}$) as a function of the ethnic community size. When the ethnic community in the host location has a size below (above) the threshold size $\widetilde{M}$, $V_h^{**} > V_h^*$ ($V_h^{**} < V_h^*$). Migrants' utility is given by eq. (19) when $M \geq \widetilde{M}$, and by eq. (17) when $M \leq \widetilde{M}$. Three different cases are shown. Diagram *(a)* represents the situation in which $\widetilde{M} < M^{**} < M^*$. This case is characterised by relatively strong preferences for ethnic products, a low degree of love for variety and/or low hostility externalities. The marginal migrant for which an ethnic sector is set up, $\widetilde{a}$ (where

$\widetilde{M} = \int_{\underline{a}}^{\overline{a}} f(a)Pda$), has a mobility cost lower than a^{**}. An intermediate case, where

$M^{**} < \widetilde{M} < M^*$, is reported in diagram *(b)*, while diagram *(c)* represents a case opposite to the first one where $M^{**} < M^* < \widetilde{M}$.

The upward sloping line in the diagrams, $C(P)$, represents mobility costs across the population in the source locality when these are uniformly distributed. The relevant equation in this simple and tractable case is:

$$C(P) = \overline{a} + P / \left(\overline{a} - \underline{a} \right) \tag{21}$$

where the intercept $\underline{a}$ is the mobility cost of the most mobile individual in the source region and the slope might be interpreted, for a given population P in the source locality, as a measure of mobility costs dispersion.

The relationship between $\widetilde{M}$ and the two optimal sizes, M^* and M^{**}, defined above is important in understanding the alternative scenarios depicted by our model and which are investigated in the next section. In **table 1,** we report the sign of the derivatives of $\widetilde{M}$, M^* and M^{**} with respect to the main variables of the model.

Table 1. The effects of model's parameters and variables on $\widetilde{M}$, M^* and M^{**} (sign of the first derivative)

	Ethnic preferences ε	Love of variety θ	Congestion costs ξ	Fixed costs of production α	Variable costs of production β	Size of native population L
$\widetilde{M}$	–	–	0	0	0	0
M^*	+	$+/-$ *	–	–	–	0
M^{**}	0	$+/-$ *	–	–	–	–

* = both signs according to the size of other parameters

As already mentioned in *section 4.3*, the threshold level $\widetilde{M}$ is decreasing in the intensity of preferences toward ethnic consumption vis-à-vis native consumption (ε) and is increasing in intensity of love for variety. The optimal size of the ethnic group when an ethnic sector already exists is increasing in ε. More migrants in the host localities expand the set of ethnic variety available to the whole ethnic community. When preferences for these goods are strong, this positive effect will dominate the negative effects over a larger size of the community. On the other hand, M^* is decreasing in congestion costs (ξ) and in labour costs of production (α, β). More ambiguous are the signs of both $\partial M^* / \partial \theta$ and $\partial M^{**} / \partial \theta$, which depend jointly on the size of other parameters in a non-linear way.

Note that an increase in the native population L, has a negative impact on M^{**}. This result is obvious from eq. (18) and is due to the fact that the bigger the native population, the greater is the number of varieties already available in the host locality. Therefore, the positive

native expansion effect associated with incoming migrants could be more than offset by negative externalities. When negative externalities are particularly strong, both M^* and M^{**} might be negative, which means that the optimal size of the ethnic community is zero.

In **figure 2**, $\widetilde{M}$, M^* and M^{**} are plotted against ε keeping all other parameters fixed at a certain level. When preferences toward type E goods are only marginally higher than those for native varieties (in the figure: $0 < \varepsilon < A$), $\widetilde{M}$ is bigger than M^* and M^{**}. This situation corresponds to diagram (c) in figure 1. Stronger ethnic preferences will lead to the alternative two cases depicted in figure 1.

4.4.3 Strategic support and the size of the ethnic community: alternative scenarios

After having outlined the general framework of our analysis, in this section we define the migration pattern under alternative scenarios. The main questions we want to address are the following:

(i)   When will established migrants provide support to newcomers?

(ii)   What is the resulting size of an ethnic group?

(iii)   Is the resulting size equal to the optimal size?

(iv)   When does an ethnic productive sector arise?

The set of possible scenarios predicted by the model is reported in **table 2**. In what follows we define the conditions under which each scenario arises and describe the underlying migration pattern.

By giving support to potential newcomers, whether sending them information on job opportunities in the host location or giving direct support in order to reduce initial difficulties

and uncertainty, established migrants influence the size of the flows.[14] Therefore, if positive externalities arise as a consequence of a larger size of the ethnic community in the host location, established migrants might strategically use 'support' as a way of internalising those externalities.

Definition 1: *The utility gain (or loss) experienced by a community of established migrants of size* M_k *as the ethnic community expands up to a size* M_j *is given by*

$$G(M_k, M_j) = M_k \left[V_{h-s}\left(M_j\right) - V_{h-s}\left(M_k\right) \right] ; [15]$$

Proposition 1: *For a given size of the ethnic community already in the host location,* M_K, *established migrants strategically support newcomers if and only if the following condition is satisfied:*

$M_k < M^*$, i.e. the size of the ethnic community is initially smaller than the optimal size M^*;

Proof: If the above condition is not satisfied and $M_k \geq M^*$, established migrants' utility is decreasing in M since $\partial V_h / \partial M \leq 0$ for any $M \in \left[M^*, P \right]$. It follows that $G\left(M_k, M_j\right) < 0$. Migrants already settled in the host locality experience a utility loss as a consequence of incoming migrants, therefore no support will be provided. ∎

When $M_k < M^*$, there exists a $M_j \in \left(M_k, M^* \right)$ such that

$$G\left(M_K, M_j\right) > \int_{M_K}^{M_j} C\left(M\right) dM - \int_{M_k}^{M_j} V_{h-s}\left(M\right) dM \geq 0 ; \text{ i.e. the gain experienced by established}$$

---

[14] And in a multi-location framework the direction of the flows.

migrants in terms of consumption externalities is larger than the support needed by a number of incoming migrants of size $\left(M_j - M_k\right)$ in order to be indifferent between migrating or remaining in the home location.

The total optimal support S^* will be:

$$S^* = \int_{M_k}^{M_{K+1}} C(M)\,dM - \int_{M_k}^{M_{K+1}} V_{h-s}(M)\,dM \tag{22}$$

where M_{K+1} is the size of the ethnic community after the new migrants triggered by the support are settled in the host location. At M_{k+1}, the marginal gain of having a new migrant is equal to the marginal support that would be necessary to induce him/her to migrate:

$$\frac{\partial G\left(M_k, M_{K+1}\right)}{\partial M} = \frac{\partial\left[C\left(M_{K+1}\right) - V_{h-s}\left(M_{k+1}\right)\right]}{\partial M}$$

Before we proceed with depicting the different scenarios reported in table 2, it is convenient to define the equilibrium size of the ethnic community as follows:

Definition 2: *The size of an ethnic community in the host location reaches a stable equilibrium (M^E) when:*

(i) $\partial G\left(M^E\right)/\partial M \leq \partial\left[C\left(M^E\right) - V_{h-s}\left(M^E\right)\right]/\partial M$ or $\partial G\left(M^E\right)/\partial M \leq 0$;

---

[15] For the autonomous initial "mass" of migrants, $M_K = M_1$ as defined by equation (16).

(ii)   for any individual with mobility cost $a_j \in \left(a^E, \bar{a} \right)$, where $M^E = \int_{\underline{a}}^{a^E} f(a) P da$, the cost

of migrating is larger than the associated net gain in utility, formally

$$C_j \left(M(a_j) \right) > V_{h-s} \left(M(a_j) \right)_j.$$

The interpretation of this definition is straightforward. A stable equilibrium, M^E, is achieved at a point in which established migrants have no incentive to further support incoming migrants and all individuals left behind in the source location have no incentive to migrate due to their high mobility costs. Condition (i) is necessary but not sufficient. A new group of individuals in the source location might find it profitable to migrate without support since the ethnic community has expanded to a level which more than compensates for their mobility costs.[16]

For the community strategic support to be effective, we must assume that established migrants act like a 'cohesive group' and are able to control and prevent free riding behaviour within the group so that all components contribute to the supportive action. In general ethnic minorities show a high degree of co-operation and mutual support, which probably is due to their ability to prevent opportunistic behaviour and enforce trust. For instance, it is frequently observed that informal financing within an ethnic minority is often a common way of obtaining credit. As emphasised by the literature, informal financing is usually the preferred option within communities where monitoring is more intense, costs of information about a debtor are very low and enforcement mechanisms are more effective.

---

[16] This will happen when the size of the ethnic community in the host destination is still lower than the optimal size ($M_K < M^*$) and therefore utility of established and incoming migrants is increasing in the size of the community, $\partial V_{h-s} / \partial M > 0$. The migrants triggered by support (*induced migrants*) are followed by a new flow of *autonomous migrants* composed by individuals who previously had incentive to stay put. As a consequence of the expansion of the ethnic community, and therefore the availability of more ethnic varieties, these individuals subsequently found it profitable to migrate even without support from established migrants.

Established migrants' support is fundamental in expanding the host locality ethnic community through subsequent *induced* migration flows. Their effort might lead to the formation of an ethnic productive sector.

Figure 3, scenario 1, illustrates the theoretical case characterised by *(i) established migrants support to newcomers, (ii) the emergence of an ethnic productive sector and (iii) an equilibrium size of the ethnic community equal to the optimal size* $M^E = M^*$. Individuals composing the initial pool of migrants, M_1, have an incentive to support newcomers since in doing so they might achieve a higher level of utility for themselves. The optimal amount of support provided is given by the total area between $C(M)$ and $V_{h-s}(M)$ as defined by eq. (22). As reported in table 2, this scenario is likely to arise when:

a) Migrants in the source locality have very strong preferences for ethnic goods and services vis-à-vis native ones. In this situation, the critical mass of migrants necessary to set up an ethnic sector is small. This is often the case for ethnic migrants with cultural, religious and social preferences very dissimilar from those of the host country. The migration cost in this case is generally higher since immigrants find it more difficult to assimilate into the mainstream society, and individuals will be willing to pay more in order to recreate the source country environment. A native language different from that of the host country will work in the same direction (see Bauer et al. 2002).

b) Mobility costs of potential incoming migrants are not prohibitively high (but neither too small, see scenario 2). The support necessary to induce incoming migrants to join the "ethnic club" in the host location should not be too large. Established migrants' support, in this case, will be positive until the ethnic community reaches the optimal size $M^E = M^*$. Above this size no support will be provided since, $\partial G(M^*)/\partial M < 0$, i.e. expansion of the community above this threshold will have detrimental effect on

established migrants' utility since the hostility externality dominates the positive consumption externality.

c) Hostility externalities or congestion costs are low.

The optimal size of the community will differ from the optimal one in **scenario 2** *(established migrants support newcomers, an ethnic productive sector is set up and the ethnic community size is $M^E \neq M^*$)* under two circumstances. Firstly, when mobility costs for those left behind in the source population are high (i.e. $C(M)$ is steep). In this circumstance, established migrants find it profitable to support up to a level where the marginal gain of having an extra migrant is equal to its marginal cost. The resulting equilibrium size of the community is smaller than the optimal, $M^E < M^*$. Secondly, the equilibrium size will differ from the optimal size also in the case of low mobility costs in the population (i.e. $C(M)$ is flat). This case is depicted in figure 3. The ethnic variety expansion effect, caused by the settlement of new migrants, will exert an attraction force also for other individuals in the host location who, therefore, benefit indirectly from established migrants' support. Joining the "migration club" becomes profitable also for them.[17]

For weak preferences toward the ethnic goods and relatively low hostility externalities, the likely outcomes are those represented in **figure 4 (scenarios 3 and 4)**. an ethnic productive sector is not set up in any of the cases but we still observe support by established migrants to newcomers. Equilibrium size of the ethnic community is $M^E = M^{**}$ in scenario 3 and $M^E \neq M^{**}$ in scenario 4.

---

[17] A strategic use of remittances in the source location might in principle prevent further migration and act as a means of achieving the optimal size. Even by including strategic remittances in the analysis, scenario (2) is still a possible outcome for very low mobility costs since the strategic remittances required to induce individuals to stay put might be too large for $M^E = M^*$ to be a feasible equilibrium.

When preferences for ethnic consumption are strong but mobility costs in the source locality are rather low, the likely outcome is depicted in **figure 5 (scenario 5)**. Migrants have a strong preference for ethnic consumption and therefore an ethnic sector is likely to emerge. Established migrants have no incentive to support newcomers, and the mobility cost curve $C(M)$ cuts the utility curve $V_{h-s}(M)$ in a point where negative externalities dominate positive ones.

Finally, no support and no ethnic productive sector will characterise the migratory experience in the presence of relatively low preferences for ethnic goods associated with strong congestion costs when migration costs are low **(figure 5, scenario 6)**. In these last two scenarios, the equilibrium size of the ethnic community is equal to the size of the initial autonomous flow of migrants $M_1 = M^E$.

5. *Concluding remarks*

In this paper we provide a theory of migration networks and ethnic communities formation based on the existence of immigrants' preferences toward ethnic non-tradable goods and services which only individuals belonging to the same ethnic community can provide. In our model, established immigrants play a key role in determining the migratory experience of the ethnic community through the strategic use of support to reduce mobility costs of subsequent migrants. The model is able to produce a rich set of scenarios, which we argue are in line with the heterogeneous migration pattern showed by different ethnic groups.

The more different the cultural, political and economic environment the region of origin is from that of the region of destination, the higher the probability that ethnic migrants will cluster together in a selected number of communities. In addition, the larger will be the resulting size of the ethnic minority since established migrants will be willing to devote more

effort to support newcomers in order to recreate, at least in part, the economic, religious and cultural environment they left behind at home.

Incoming migrants expand the set of ethnic non-tradable varieties which are produced in the host locality. This effect is undoubtedly positive on migrants already settled in the ethnic cluster. Nevertheless, migrants compete over scarce resources such as housing and job opportunities. In addition, as human history shows, as ethnic minority expands in the host destination the probability of hostile and discriminatory behaviour toward them from the native population increases as well. In our model we show that coexistence of these positive and negative externalities determines an "optimal size" of the ethnic community. The actual size of the community may well be different from the optimal size.

The model's predictions match some important findings observed in empirical studies. When a group of migrants has similar preferences to those of the native population in the host country, our model predicts more dispersed migration flows that are mainly driven by regional differences in economic fundamentals (real wages, unemployment rate, etc.). This should be the case for internal migration which involves individuals characterised by relatively homogeneous preferences. In this regard, the model is in line with the different geographical patterns of internal versus international migration flows frequently found in empirical studies. Immigrants are generally more geographically concentrated than natives and reside in cities with a large population of the same ethnic group (see for instance Bartel 1989).

Often ethnic clusters tend to die with the ageing of the first generation of established migrants. On the basis of the premises of our model this phenomenon could be explained by decreasing attachment to "ethnic consumption" and increasing assimilation into native society by descendants of those individuals.

Interesting extensions of the model and alternative formulation are possible. First, it would be interesting to extend the model by considering a multi-location framework. The advantage would be one of analysing not only the size of the ethnic community but also the direction of migration in a richer set. Once a "migration club" has eroded its attractiveness, new destinations might emerge and the migratory pattern is replicated in the new destination. Second, considering more explicitly the welfare of native population and endogenising the congestion/hostility externalities could considerably enrich the migration pattern described above.

Table 2 – Migration pattern: features of alternative scenarios

		Strategic support by established migrants		
		Yes		**No**
		Optimal size of the ethnic community $\left(M^*\right)$		**Scenario (5)**
		Yes	**No**	
		Scenario (1)	**Scenario (2)**	
Ethnic Sector	Yes	▪ Strong preferences for ethnic goods and service $(\varepsilon \to high)$; ▪ Large dispersion of mobility costs in the source location population (slope of $C(M)$, i.e. $1/\left(\bar{a}-\underline{a}\right) \to relatively\ small$); ▪ Low hostility externalities $(\xi \to low)$.	▪ Strong preferences for ethnic goods and service $(\varepsilon \to high)$; ▪ Very large (very small) dispersion of mobility costs in the source location population (slope of $C(M)$, i.e. $1/\left(\bar{a}-\underline{a}\right) \to$ either very large or very small); ▪ Low hostility externalities $(\xi \to low)$.	▪ Strong preferences for ethnic goods and service $(\varepsilon \to high)$; ▪ Small dispersion of mobility costs in the source location population (slope of $C(M)$, i.e. $1/\left(\bar{a}-\underline{a}\right) \to large)$; ▪ High hostility externalities $(\xi \to low)$.
		Optimal size of the ethnic community $\left(M^{**}\right)$		**Scenario (6)**
		Yes	**No**	
		Scenario (3)	**Scenario (4)**	
	No	▪ Weak preferences for ethnic goods and service $(\varepsilon \to 0)$; ▪ Large dispersion of mobility costs in the source location population. (slope of $C(M)$, i.e. $1/\left(\bar{a}-\underline{a}\right) \to relatively\ small$); ▪ Low hostility externalities $(\xi \to low)$.	▪ Weak preferences for ethnic goods and service $(\varepsilon \to 0)$; ▪ Very large (very small) dispersion of mobility cost in the source location population (slope of $C(M)$, i.e. $1/\left(\bar{a}-\underline{a}\right) \to$ either very large or very small); ▪ Low hostility externalities $(\xi \to low)$.	▪ Weak preferences for ethnic goods and service $(\varepsilon > 0)$; ▪ Small dispersion of mobility costs in the source location population (slope of $C(M)$, i.e. $1/\left(\bar{a}-\underline{a}\right) \to large)$; ▪ High hostility externalities $(\xi \to low)$.

References:

Banerjee B. (1984), Information Flow, Expectations and Job Search, *Journal of Development Economics,* Vol. 15, pp. 239-257

Bartel A. P. (1989), Where Do the New U.S. Immigrants Live?, *Journal of Labor Economics,* Vol. 7 (4), pp. 371-391

Bauer T., G. S. Epstein and I. Gang (2002), Enclaves, Language and the Location Choice of Migrants, *CEPR Discussion Paper Series*, n. 3527, September 2002

Bauer T., G. S. Epstein and I. Gang (2002), Herd Effects or Migration Networks? The location Choice of Mexican Immigrants in the US, *CEPR Discussion Paper Series*, n. 3505, August 2002

Blalock H. M. (1967), Toward a Theory of Minority-Group Relations, Wiley, New York.

Borjas G. J. (1995), Ethnicity, Neighborhoods, and Human-Capital Externalities, *American Economic Review*, Vol. 85 (3), pp. 365-390

Boyd R. L. (1998), The Storefront Church Ministry in African American Communities of the Urban North During the Great Migration: The Making of an Ethnic Niche, *The Social Science Journal*, Vol. 35 (3), pp. 319-332

Carlton D. W. (1995), Economic Organization and Conflicts, *Journal of Institutional and Theoretical Economics*, Vol. 151 (1), pp. 218-236

Carrington W., Detragiache E. and Vishwanath T. (1996), Migration with Endogenous Moving Costs, *American Economic Review*, 86 (4), pp. 909-930

Chau N. H. (1997), The Pattern of Migration with Variable Migration Cost, *Journal of Regional Science*, Vol. 37 (1), pp. 35-54

Chiswick Barry R. and Paul W. Miller (2003), Do Enclaves Matter in Immigrants Adjustment?, IZA Discussion Paper No. 449

Chiswick Barry R., Lee Y. L. and Paul W. Miller (2002), The Determinants of the Geographic Concentration among Immigrants: Application to Australia, IZA Discussion Paper No. 462

Docquier F. and H. Rapoport (1998), Are Migrant Minorities Strategically self-selected?, *Journal of Population Economics*, Vol. (11), pp. 579-588

Epstein G. S. (2002), Informational Cascades and the Decision to Migrate, IZA Working Paper n. 445

Glazer A., Gradstein M. and P. Ranjan (2003), Consumption variety and urban agglomeration, *Regional Science and Urban Economics*, Vol. 33 (1), pp. 653-661

Glebe G. (1986), Segregation and intra-urban mobility of a high-status ethnic group: the case of Japanese in Dusseldorf, *Ethnic and Racial Studies*, Vol. (9), pp. 461-483

Gonzales A. (1998), Mexican Enclaves and the Price of Culture, *Journal of Urban Economics*, vol. 43, pp.273-291

Grannovetter M. (1973), The Strength of Weak Ties, *American Journal of Sociology*, Vol. 78, pp. 1360-1388

Grannovetter M. (1982), The Strength of Weak Ties: a network theory revissited, in *Social Structure and Network Analysis*, Peter V. Marsden and Nan Lin eds., pp. 105-130, Sage Publication (Beverly Hills)

Grossman J. R. (1989), Land of Hope: Chicago, Black Southerners, and the Great Migration, Chicago, University of Chicago Press.

Kosa J. (1956), Hungarian Immigrants in North America: their Residential Mobility and Ecology, *The Canadian Journal of Economics and Political Science*, Vol. 22 (3), pp. 358-370

Lever-Tracy C. and R. Holton (2001), Social exchange, reciprocity and amoral familism: aspects of Italian chain migration to Australia, *Journal of Ethnic and Migration Studies*, Vol. 27, n.1, pp.81-99

Munshi K. (2003), Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market, *Quarterly Journal of Economics*, Vol. 118 (2), pp. 549-599

Nelson P. (1959), Migration, real income and information, *Journal of Regional Science*, Vol. (1), pp. 43-74

Rapoport H. and A. Weiss (2003), The Optimal Size for a Minority, *Journal of Economic Behaviour and Organization*, Vol. 52 (1), pp. 27-45

Saxenian A. (1999), Silicon Valley's New Immigrant Entrepreneurs, Public Policy Institute of California, San Francisco (CA)

Stark O. (1999), Altruism and Beyond: An Economic Analysis of Transfers and Exchange within Families and Groups, Cambridge University Press, Cambridge

Stark O. and Wang Y. Q. (2002), Migration dynamics, *Economics Letters*, Vol. 76 (2), pp. 159-164

Taylor J. E. (1986), Differential Migration, Networks, Information and Risk, in *Migration, Human Capital and Development*, ed. Oded Stark, JAI Press Inc., Greenwich Connecticut.

Taylor M. (1998), Local Racial/Ethnic Proportions and White Attitudes: Number Count, *American Sociological Review*, Vol. 63, pp. 56-78

White P. (1998), The Settlement Patterns of Developed World Migrants in London, *Urban Studies*, Vol. 35 (10), pp. 1725-1744

Wilson T. D. (1998), Weak Ties, Strong Ties: Network Principles in Mexican Migration, *Human Organization*, Vol. 57 (4), pp. 394-403

Winters P., de Janvry A. and E. Sadoulet (2001), Family and Community Networks in Mexico-U.S. Migration, *Journal of Human Resources*, Vol. 36 (1), pp. 159-184

Zavodny M. (1999), Determinants of recent Immigrants' Locational Choices, *International Migration Review*, Vol. 33, pp. 1014-1030

Appendix I:

In the destination location, given a certain wage level (w), the migrant maximises utility (1) subject to the budget constraint (2), we set up the Lagrangian:

$$L = \int_{i=0}^{n_E} ((1+\varepsilon)c_{Ei})^\theta \, di + \int_{j=0}^{n_Y} ((1+\varepsilon)c_{Yj})^\theta \, dj - \xi M + \lambda \left[w - \int_{i=0}^{n_E} p_{Ei}c_{Ei} \, di - \int_{i=0}^{n_E} p_{Yj}c_{Yj} \, dj \right] \qquad \text{(A.1)}$$

the F.O.Cs. for (A.1) are the following:

$$\partial L / \partial c_{Ei} = \theta((1+\varepsilon)c_{Ei})^{\theta-1}(1+\varepsilon) - \lambda p_{Ei} = 0 \qquad \text{(A.1.1)}$$

$$\partial L / \partial c_{Yj} = \theta(c_{Yj})^{\theta-1} - \lambda p_{Yj} = 0 \qquad \text{(A.1.2)}$$

$$\partial L / \partial \lambda = w - \int_{i=0}^{n_E} p_{Ei}c_{Ei} \, di - \int_{i=0}^{n_E} p_{Yj}c_{Yj} \, dj = 0 \qquad \text{(A.1.3)}$$

re-arranging (A.1.1)- (A.1.3) we get:

$$c_{Ei} = \left(\frac{\theta}{p_{Ei}} \right)^{\frac{1}{1-\theta}} (1+\varepsilon)^{\frac{\theta}{1-\theta}} ; \quad c_{Yj} = \left(\frac{\theta}{p_{Yj}} \right)^{\frac{1}{1-\theta}} \qquad \text{(A.1.4)}$$

Given that E and Y varieties are produced using the same IRS technology, $p_{Ei} = p_{Yj}$ and defining as in Glazer et al (2003) $\varepsilon' = (1/(1+\varepsilon))^{\theta/(1-\theta)}$, from (A.1.4) it follows:

166

$$c_{Yj} = c_{Ei}\varepsilon'$$

(A.1.5)

It is worth noting that $0 \le \varepsilon' \le 1$, therefore eq. (A.1.5) indicates that the amount consumed of a Y good is only a fraction of that of an E good; this is obvious given that migrants attach more value to non-tradable ethnic varieties. Besides, as love-of-variety (θ) and migrants' preference for ethnic variety (ε) increase, migrants will shift consumption from native to ethnic varieties ($\partial \varepsilon'/\partial \varepsilon < 0$ and $\partial \varepsilon'/\partial \varepsilon < 0$).

Using the budget constraint and (A.1.5) we can write the migrants' demand for a single E and Y variety as follows:

$$c_{Ei} = \frac{w}{n_E p_E + n_Y p_Y \varepsilon'} \quad ; \quad c_{Yj} = \frac{w\varepsilon'}{n_E p_E + n_Y p_Y \varepsilon'}$$

(A.1.6)

Appendix II:

Using (3), (4) and (6) into (1) and (9), $V^{no\_ethnic} = V$:

$$n_Y^{no\_ethnic}\left(\frac{w\theta}{n_Y^{no\_ethnic}\beta w}\right)^\theta - \xi M =$$

$$= n_E\left((1+\varepsilon)\frac{w\theta}{n_E \beta w + n_Y \beta w\varepsilon'}\right)^\theta + n_Y\left(\frac{w\theta\varepsilon'}{n_E \beta w + n_Y \beta w\varepsilon'}\right)^\theta - \xi M$$

substituting (11) and (13) into this last expression yields:

$$
(M+L)\left(\frac{1-\theta}{\alpha}\right)\left[\frac{\theta}{(M+L)\left(\frac{1-\theta}{\alpha}\right)\beta}\right]^{\theta} =
$$

$$
=\left(M-\frac{L\varepsilon'}{1-\varepsilon'}\right)\left(\frac{1-\theta}{\alpha}\right)\left[(1+\varepsilon)\frac{\theta}{\left(M-\frac{L\varepsilon'}{1-\varepsilon'}\right)\left(\frac{1-\theta}{\alpha}\right)\beta+\left[L/(1-\varepsilon')\right]\left(\frac{1-\theta}{\alpha}\right)\beta\varepsilon'}\right]^{\theta} +
$$

$$
+\left(\frac{L}{1-\varepsilon'}\right)\left(\frac{1-\theta}{\alpha}\right)\left[\frac{\theta\varepsilon'}{\left(M-\frac{L\varepsilon'}{1-\varepsilon'}\right)\left(\frac{1-\theta}{\alpha}\right)\beta+\left[L/(1-\varepsilon')\right]\left(\frac{1-\theta}{\alpha}\right)\beta\varepsilon'}\right]^{\theta}
$$

If we substitute in the above expression, **condition 1** $\rightarrow$ $\widetilde{M}=\dfrac{L\varepsilon'}{1-\varepsilon'}$, we obtain:

$$
\left(\frac{L}{1-\varepsilon'}\right)\left(\frac{1-\theta}{\alpha}\right)\left[\frac{\theta}{\left(\frac{L}{1-\varepsilon'}\right)\left(\frac{1-\theta}{\alpha}\right)\beta}\right]^{\theta} = \left(\frac{L}{1-\varepsilon'}\right)\left(\frac{1-\theta}{\alpha}\right)\left[\frac{\theta}{+\left(\frac{L}{1-\varepsilon'}\right)\left(\frac{1-\theta}{\alpha}\right)\beta}\right]^{\theta}
$$

$V^{no\_ethnic}=V$

Appendix III:

Comparing equations (18) and (20) it is easy to check that M^* is always larger than M^{**}:

$$
M^*-M^{**}=\frac{\theta}{\beta}(1-\theta)^{\frac{2-\theta}{\theta}}\alpha^{-\frac{1}{\theta}}\xi^{\frac{1}{\theta}}\left[\varepsilon+L\right]>0 \tag{A.3.1}
$$

Analogously it can be shown that, for any $M > \underline{M} = L\dfrac{\varepsilon'}{1-\varepsilon'}$, utility in *case 2* is always

higher than in *case 1*. Using (17) and (19) and given that there exists a $\psi \in]0,1[$ for which

$M\psi = \underline{M} = L\dfrac{\varepsilon'}{1-\varepsilon'}$ we have:

$$
\begin{aligned}
V_M^y - V_M^{yy} &= \left(\frac{\theta}{\beta}\right)^{\theta}\left(\frac{1-\theta}{\alpha}\right)\left[M^{1-\theta}(1+\varepsilon)^{\theta}-(M+L)^{1-\theta}\right] = \\
&= \left(\frac{\theta}{\beta}\right)^{\theta}\left(\frac{1-\theta}{\alpha}\right)\left[M^{1-\theta}(1+\varepsilon)^{\theta}-\left(M+\frac{M\psi(1-\varepsilon')}{\varepsilon'}\right)^{1-\theta}\right] = \qquad (A.3.2) \\
&= \left(\frac{\theta}{\beta}\right)^{\theta}\left(\frac{1-\theta}{\alpha}\right)M^{1-\theta}(1+\varepsilon)^{\theta}\left[1-\left(\varepsilon'+\psi(1-\varepsilon')\right)^{1-\theta}\right] > 0
\end{aligned}
$$

where the last term in square brackets is always positive (since $0 < \varepsilon' < 1$, $\forall \varepsilon > 0$ and

$\forall \theta \in]0,1[$).

Figure 1 - Migrants' indirect utility as a function of the ethnic community size
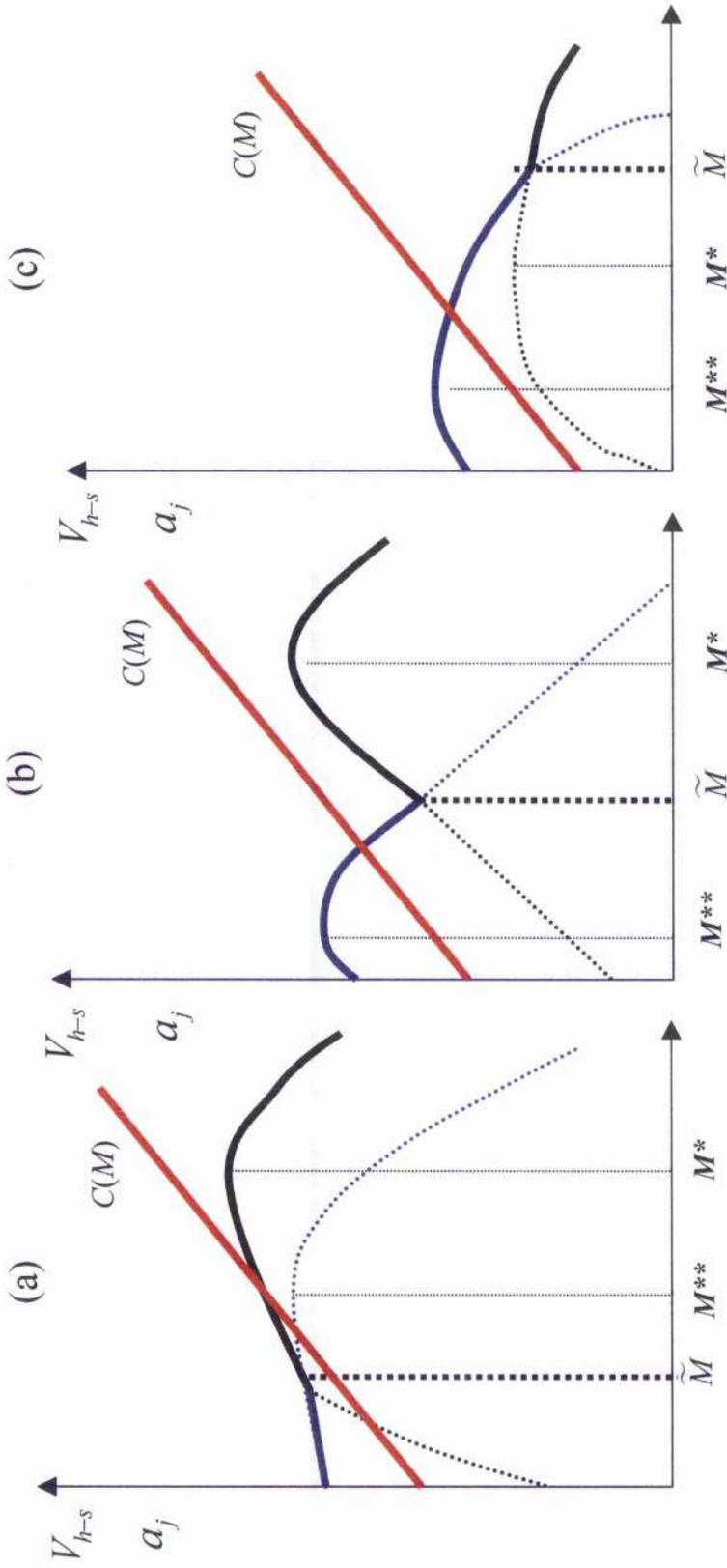
Figure 2 - Intensity of the preferences towards the ethnic varieties and $\widetilde{M}, M^*, M^{**}$

M^*

M^{**}

$\widetilde{M}$

ε

M
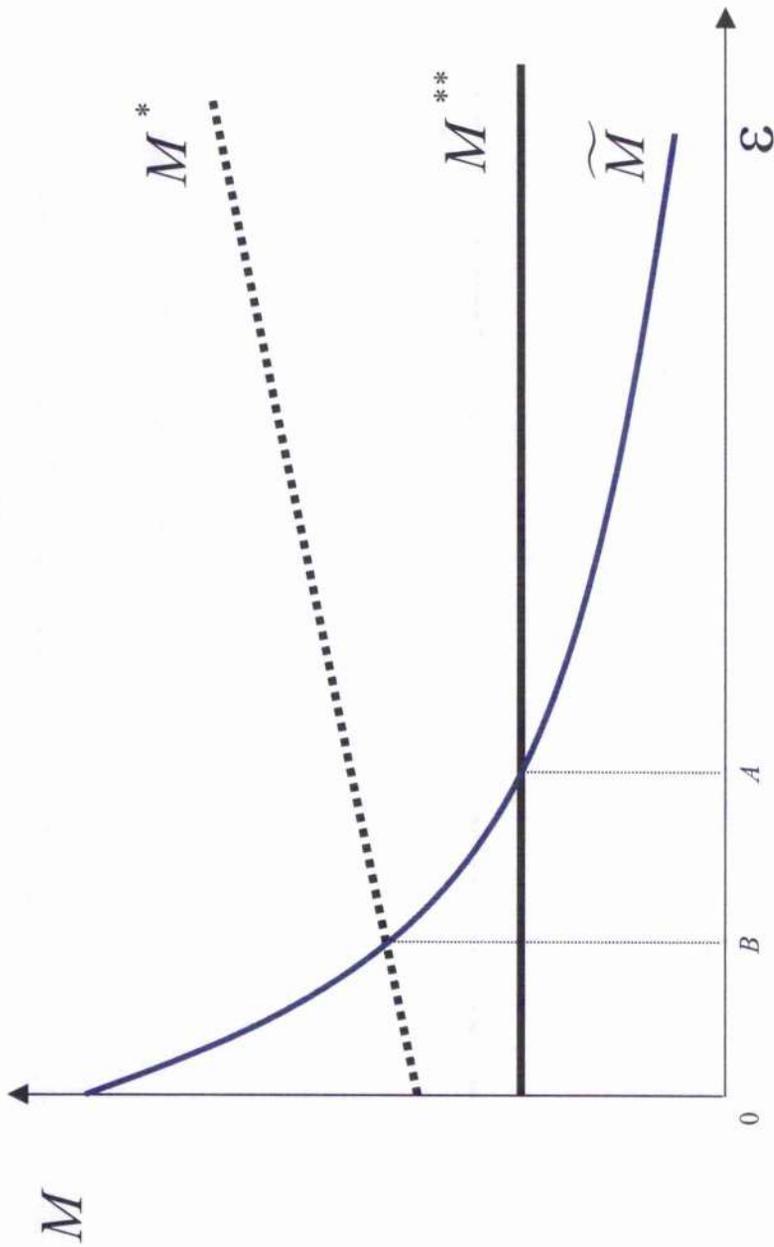
$0 < \varepsilon < A \Rightarrow \left(M^{**} < M^* < \widetilde{M} \right)$ Figure 1- diagram (c)

$A < \varepsilon < B \Rightarrow \left(M^{**} < \widetilde{M} < M^* \right)$ Figure 1- diagram (b)

$\varepsilon > B \Rightarrow \left(\widetilde{M} < M^{**} < M^* \right)$ Figure 1- diagram (a)

Figure 3

scenario (1)
ethnic sector, support & $M^E = M^*$

scenario (2)
ethnic sector, support & $M^E > M^*$

$G(M_k, M^E) > S(M_k, M^E)$
and
$\partial G(M_k, M^E) / \partial M = 0$

$G(M_1, M_1) > S(M_1, M_1)$

V_{h-s}

a

M^{**} $M_1 \widetilde{M}$ $M^E = M^*$

M^{**} M_1 $\widetilde{M}$ M^{**} M^E

172
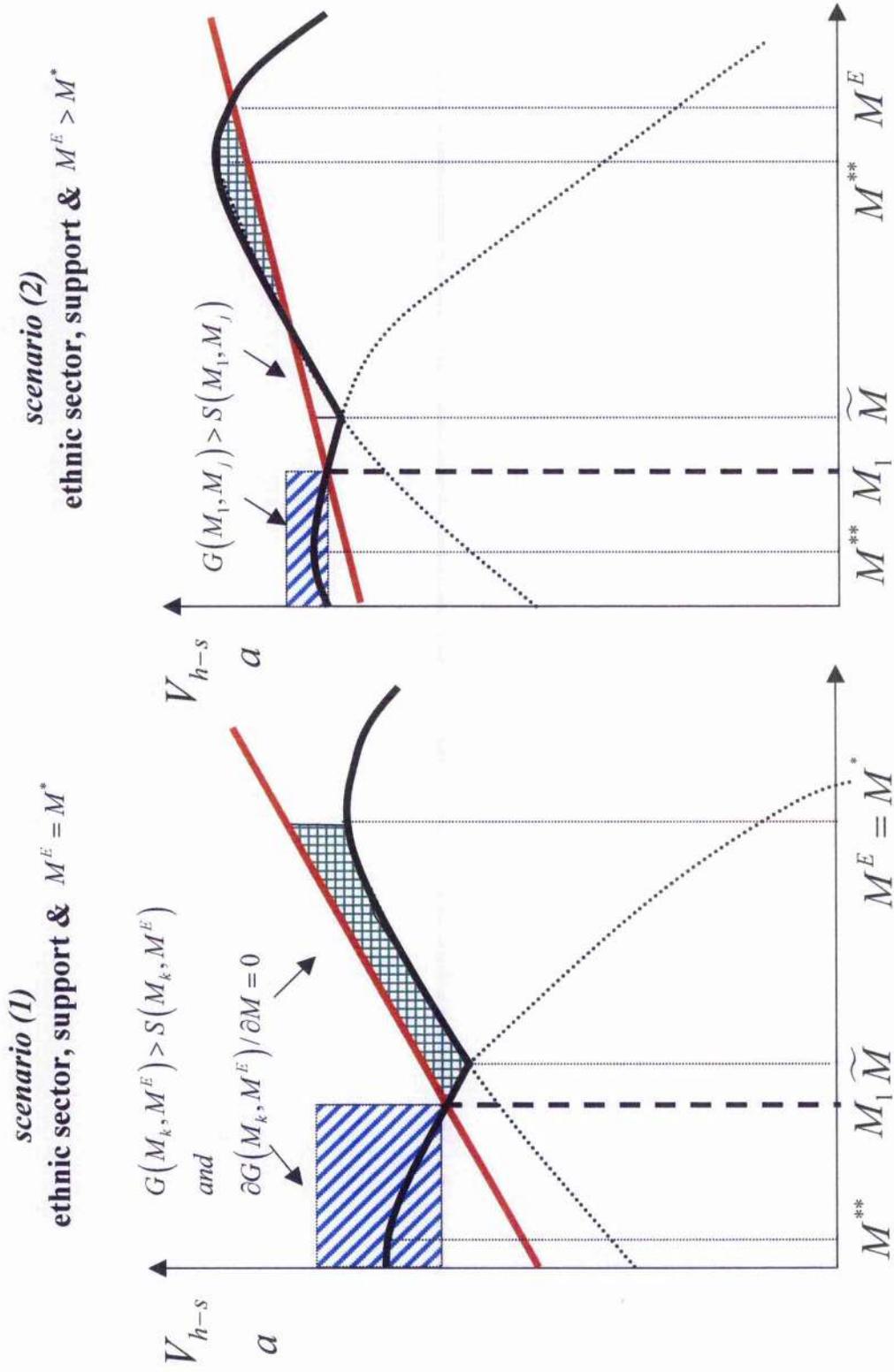
Figure 4

scenario (3)
no ethnic sector, support & $M^E = M^{**}$
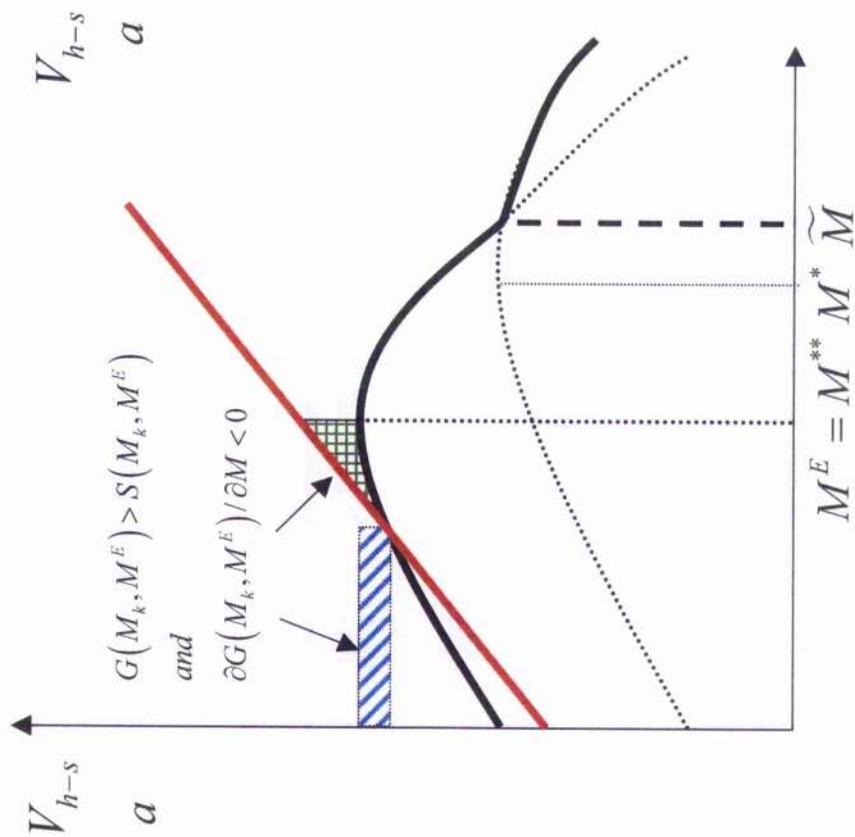
*scenario (4)*
no ethnic sector, support & $M^E < M^{**}$

$$G(M_k, M^E) > S(M_k, M^E)$$
$$and$$
$$\partial G(M_k, M^E) / \partial M < 0$$

$$\partial G(M_k, M^E) / \partial M = \partial S(M_k, M^E) / \partial M$$

V_{h-s}
a

$M^E = M^{**} M^* \widetilde{M}$

$M^E \quad M^{**} \quad M^* \quad \widetilde{M}$

Figure 5

scenario (5)
ethnic sector and no support

scenario (6)
no ethnic sector and no support

V_{h-s}

a

V_{h-s}

a

$\widetilde{M}$ M^* $M^E = M_1$

M^* $M^E = M_1$ $\widetilde{M}$

note: $M^{**} < 0$

174