NBER WORKING PAPER SERIES

THE SELECTION OF MIGRANTS AND RETURNEES: EVIDENCE FROM ROMANIA AND IMPLICATIONS

J. William Ambrosini
Karin Mayr
Giovanni Peri
Dragos Radu

Working Paper 16912
http://www.nber.org/papers/w16912

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2011

The authors thank Frederic Docquier, Christian Dustmann, Giovanni Facchini and David Newhouse for very helpful comments and suggestions. The authors are most indebted to the governments of Germany, Austria, Norway, Korea, and Switzerland for their financial support through the Multi-Donor Trust Fund for "Labor Markets, Job Creation, and Economic Growth". The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Selection of Migrants and Returnees: Evidence from Romania and Implications
J. William Ambrosini, Karin Mayr, Giovanni Peri, and Dragos Radu
NBER Working Paper No. 16912
March 2011
JEL No. F22,J61,O15

ABSTRACT

This paper uses census and survey data to identify the wage earning ability and the selectivity of recent Romanian migrants and returnees. We construct measures of selection across skill groups and estimate the average and the skills-specific premium for migration and return for three typical destinations of Romanian migrants after 1990. We find evidence for a sorting of migrants consistent with skill compensation in destination countries. The premium to return migration increases with migrants' skills and drives the positive selection of returnees. Based on the rationality of these migration decisions, a model of education, migration and return predicts positive long-run effects of increased migration for average skills and wages in Romania.

J. William Ambrosini
University of California at Davis
Department of Economics
One Shields Avenue
Davis, CA 95616
will@ambrosini.us

Karin Mayr
Department of Economics
University of Vienna
Vienna, Austria
karin.mayr@univie.ac.at

Giovanni Peri
Department of Economics
University of California, Davis
One Shields Avenue
Davis, CA 95616
and NBER
gperi@ucdavis.edu

Dragos Radu
Policy Studies Institute
50 Hanson Street,
London, W1W 6UP
United Kingdom
D.Radu@psi.org.uk

1 Introduction

The emigration of the most productive workers, attracted by higher wages abroad, has often been regarded as harmful for developing countries (Gruber and Scott 1966; Bhagwati 1976; Bhagwati and Hamada 1974; Bhagwati and Rodriguez 1975). However, from the perspective of migrants themselves, migration is an opportunity to improve, sometimes dramatically, their standard of living. There is evidence that migrants e.g. from Latin America (Clemens, Montenegro and Pritchett 2008), from India (De Coulon and Wadsworth 2010) or from Eastern Europe (Budnik 2009) earn on average two to three times more at destination than they would at home. Moreover, the migration of highly skilled may induce virtuous educational incentives in the native population. In the long run, this might increase the overall human capital of the sending country. "Brain gain" has been identified as a theoretical possibility[1] and recent empirical research has looked for evidence of it in the data. Beine et al. (2001, 2008) use a cross-country approach to show that low emigration rates are positively correlated with average schooling levels. Using individual data, Batista et al. (2007) and Chand and Clemens (2008) find a positive incentive effect of skilled emigration on education. In addition to this, a large body of work has shown migration tends to be temporary rather than permanent and return migrants often become successful entrepreneurs or return with highly productive skills. The positive impacts of return migration for the countries of origin have been analyzed theoretically by Dustmann (1995), Santos and Postel-Vinay (2003), Mayer and Peri (2009), Dustman, Fadlon and Wiess (2010) and Dustmann and Glitz (2011), among others. There is extensive evidence that return migrants receive income premia for their work experience abroad (Reinhold and Thom 2009; Barret and Goggin 2010). Several recent studies have also emphasized the importance of returnees as a source of entrepreneurship (Constant and Massey 2002; McCormick and Wahba 2001).[2]

These aspects of migration and return are particularly relevant for the case of Central and Eastern Europe (CEE). After the opening of the borders in 1990 and subsequently in the context of EU-enlargement, several of its professionals as well as unskilled workers moved to Western Europe and to North America (see e.g. Kahanec and Zimmermann 2009 for a recent overview). Over the last two decades, return migrants also became an important and fast-growing group on the labour markets in all CEE (for a recent overview see e.g. Martin and Radu 2011). Precise and comparable figures of the stock of East European migrants who already returned to their countries of origin are still missing. However, some recent research suggests that these migrants acquire productive skills while abroad and receive significant income premia upon return (see e.g. Co, Gang and Yun 2000 for female return migrants in Hungary; De Coulon and Piracha 2005 for Albanian returnees; Hazans 2008 for Latvian returnees and Iara 2009, Martin and Radu 2011 for cross-country comparisons). There is also evidence that returnees in CEE have a higher propensity for entrepreneurial activities or for self-employment

[1] For instance in the papers by Stark et al. (1997), (1998), Mountford (1997), and Beine, Docquier and Rapoport (2001).

[2] Returnees have been sources of start-ups in high-tech sectors in countries such as India (Commander et al. 2008) and in the Hsinchu Science Park in Taipei (Luo and Wang 2002). Zucker and Darby (2007) find that in the period 1981-2004 there was a strong tendency of "star scientists" in several fields in the US to return to their country of origin and promote the start-up of high-tech firms (especially in China, Taiwan and Brazil).

than non-migrants (Kilic et al. 2009; Piracha and Vadean 2010), but are at the same time more likely not to participate on the labour market.

The consequences of migration and return on the sending countries will depend crucially on two aspects: the size and the selection of these flows. The larger the number of migrants and returnees the larger are the potentials for gains and losses. Moreover, for the countries of origin, a positive selection of migrants and returnees, in terms of their skills, may represent both a challenge (risk of brain drain) and an opportunity (incentives for learning and improvement of skills). Given the lowering barriers to international mobility from CEE, will increased mobility of Eastern Europeans result in harmful consequences for their countries of origin? How does migration and return contribute to the productivity and income of workers? What will be the consequences of further reducing the cost of migration?

The aim of our paper is to quantify the size and selection of migration and return for the case of a representative CEE country and to characterize the potential consequences of international mobility for its levels of wages and productive skills.

Our empirical analysis will focus on migration from Romania. This is a particularly relevant case due to the size of the population, but also the structure and destination of labour outflows and the considerable rate of return migration, which are average among CEE countries.

Using a unique combination of census and large survey data we are able to identify Romanian migrants in three main destination countries (Spain, Austria and the US) and to match this information with micro-data on non-migrants and returnees in Romania. We use census data (2000-2001) from these destination countries as well as data from the EU-Survey on Income and Living Conditions (EU-SILC) for Austria and Spain, and from the National Demographic Survey (NDS) of Romania (2003). These data provide a picture of the relative size and relative characteristics (including wages) of the cross-section (circa 2002) of individuals from Romania who are non-migrants, returnees, migrants to the US, Austria or Spain. These three countries were typical destination countries for Romanian and other CEE migrants but also span very well the different ranges of institutions and labor market types across which CEE migrants sorted after 1990.

Our results suggest that migration choices are responsive to economic incentives as workers in specific skill cells migrate in larger shares to countries which pay higher wage premia for those skill cells. We observe that migrants to the US are positively selected because the wage premium of migrating to the US is much higher for the high skill-cells (in terms of wage earning ability). To the contrary migrants to Spain are more likely to come from low-skill cells as the wage premium of migrating to Spain is larger for low skills cells. Austria exhibits a migration premium neutral to skill level. This rationality of migration is consistent with other findings for CEE migrants (e.g. Budnik 2009). Romanian returnees are positively selected and this also supports the other finding of higher return premium for highly skilled. Given our data we can only identify selection on observable characteristics, and we refer to the literature to argue how selection on unobserved characteristics, though unlikely to reverse any of them, may affect the results.

The rationality of migration and the positive selection of returnees allows us to evaluate the aggregate (skill and wage) effects of migration for Romania. In order to quantify these effects, we use the estimated returnee premium and the observed scale of return-migration in a model of schooling, migration and return, developed previously by Mayr and Peri (2009). We adapt the parameters to the case of Romania to obtain the long-run impact of increased mobility, accounting both for return migration and for indirect effects from incentives on schooling.

The rest of the paper is organized as follows. Section 2 presents some stylized facts and aggregate statistics on migration and return for CEE and Romania. Section 3 describes our data and the measures of average selection and average premium we construct. Section 4 presents our estimates for selection and return premia. Section 5 shows empirical evidence of the relation between migration frequency and premia across skill groups. Section 6 uses some of our estimates and summary statistics in a model simulating the skill and wage effect of relaxing migration constraints for Romania. Section 7 concludes the paper.

2 Stylized facts of East European Migration

2.1 Stocks of Migrants and Returnees

Prima facie evidence on migration and return can be obtained from population censuses (or current population surveys). For years around 2000 such data have been organized by Docquier and Marfouk (2006) into widely used data set. For a selected group of receiving OECD countries, the United Nations (2009) has also collected long time series on yearly gross inflows of migrants by country of origin.

Although these data are fraught with various methodological problems, they provide a first take on the magnitude of migration from and return to Eastern Europe. We consider the years 1990 and 2000 for a group of 14 Eastern European countries that can be identified consistently. Table 1 shows the stock of emigrants to all OECD countries from each of these Eastern European countries, as shares of their domestic population. As expected, every single country increased, sometimes very significantly, the share of its migrant population abroad. Around the year 2000 these shares varied between 3-4% in large countries (e.g. Poland and Romania) and 17-19% in smaller countries (e.g. Albania and Macedonia).

For a subset of these Eastern European sending countries and a subset of OECD receiving countries we can also impute the return migrants as share of the gross flows. We use the data on stocks from Docquier and Marfouk (2006) to obtain (by difference) the net immigration country by country between 1990 and 2000. From the United Nation (2009) data on yearly gross flows from the same countries of origin to the same destinations we obtain the cumulative gross flows of migrants (1991-2000). The difference between gross flows (from country i to j) and the net changes of people from country i living in country j constitutes a measure of re-migration[3].

[3] We ignore the mortality of migrants.

Following Borjas and Bratsberg (1996) and Dustmann and Weiss (2007) we assume that all of these re-migrants are returnees and we count this difference as the measure of imputed returnees. We summarize the results in Table 2 and 3 by aggregating gross and imputed return flows by source and host country (respectively). Table 2 is most useful for understanding the import of return migration in a period that experienced large gross migration from Eastern Europe. Three features are worth of notice.

First, and most importantly, for all countries of origin and destination the imputed return migration is a substantial share of total gross migration flows. The ratio of returnees to gross migrants can be larger than 1 since not only migrants who arrived in this decade but also earlier migrants returned during this period. Table 1 shows that Czechoslovakia, Hungary and Poland experienced return migration close or even larger than their gross emigration flows. This is an expected finding since these countries have a longer history of migration before 1989 and might experience now a "retirement migration" of migrants who left around 1968. However, all other East European countries have return rates between 0.3 and 0.6. This suggests a substantial role of return migration on the labour market in countries of origin. Second, the summary statistics by destination countries (Table 3) show that most countries display rates of return migration to Eastern Europe which are considerable (for Australia and France exceptionally high) but consistent with previous reports of the retention rates of migrants in OECD destinations (OECD 2008, part III). Third, Table 2 also shows that Romania is an average East-European country with regard to the rate of return migration relative to gross flows. The median return rate for the considered countries is 1.12 returnees per 2 migrants, while for Romania it is slightly below 1 returnee per 2 migrants.

These aggregate figures have to be interpreted with caution. They may be biased if, for instance, undocumented migrants are better counted in census data than in the official entry statistics, or if the definition of immigrants (by nationality, place of birth or country of last residence) is not consistent between census and administrative data. Despite such measurement problems we can certainly conclude that return migration is not a marginal phenomenon for East European countries. We can safely say that a substantial fraction of migrants eventually return to their home country. Assuming the return rate was not decreasing in the decades before 2000, between 30 and 60% of emigrants from Eastern Europe eventually return to their home country. To grasp the likely impact of this migration on the home labour markets it is important to understand the rationality of migrants and returnees. There are extremely diverse reasons for return migration, ranging from the completion of a studies abroad or the termination of fixed term temporary contracts to improved opportunities at home. However, we aim to characterize the skills of migrants and returnees to determine if the behavior of Romanian migrants is consistent with economic rationality in which changing labor market opportunities produce incentives for migration.

2.2 Romanian Migration Patterns in the early 2000s

In the period immediately after the regime change and the opening of the borders in 1989, migration from Romania was first characterized by mass permanent emigration of ethnic minorities (German and Hungarian). However, by the mid 1990s a new pattern of labour migration to various European and overseas destinations emerged. Labour outflows increased steadily against the background of a slow pace of economic restructuring which resulted in a large decline in GDP, high inflation, mass layoffs, decreasing real wages and rising unemployment (Earle and Pauna 1996, 1998). Deindustrialization led to a decrease of industrial employment by almost 3 million jobs and affected both younger and older workers, who, because they lacked skills or did not have the incentive to acquire new skills, were less likely to find new employment opportunities (Voicu, 2002).

During the period 1994-2000, while overall flows were still rising, the destinations of Romanian migrants became more settled. With regard to the selectivity of the migration flows from Romania, these destination countries could be grouped into three main categories (EIR 2003, Diminescu and Lazaroiu 2002, Radu 2003, Baldwin-Edwards 2007). First, positive selection was observed for migration flows to more traditional immigration countries (US, Canada, Australia). These flows were rather small but persistent and included a significant share of young people who migrate for educational purposes (Diminescu 2003). In the early 2000s the US was one of the main countries from where migrants returned and settled back in Romania (OECD 2008). A second group of destination countries were characterized by a neutral average selectivity of migration from Romania often determined by a polarization of migration flows towards the lower and the upper ends of the skill distribution. These were the continental European countries which, despite a "Fortress Europe" migration policy, received most of the East European migrants over the 1990s: Germany, Austria and France (Sandu et al. 2006; Diminescu 2003). Third, particularly towards the end of the 1990s and the early 2000s large flows of Romanian migrants arrived in Mediterranean countries, mainly in Spain and Italy, but also, to a lesser extent, in Portugal and Greece (Sandu et al. 2006, Diminescu and Lazaroiu 2002). These flows were characterized by negative selection: most migrants were lower skilled, tended to have a longer migration history, often involving informal or illegal employment spells, and made use of network ties established in their communities of origin (Elrick and Ciobanu 2009; Şerban and Voicu 2010).

This pattern of migration consolidated in the period 1994-2001. After that, the Schengen visa restrictions were lifted for Romanian citizens. The numbers continued to rise, particularly to the Southern European destinations, but qualitatively the pattern remained the same (Sandu et al. 2006).

For the period 2002-2003 we can construct a measure of the stock of Romanian migrants in OECD countries and a measure of returnees, both as shares of the total population in Romania. Moreover, we can also characterize the distribution of migrants residing in OECD countries (using data from Docquier and Marfouk 2006) and of returnees (using microdata from the NDS 2003) by education. These data are summarized in Table 4 and confirm that the group of returnees as of 2003 was slightly larger than the group of Romanians abroad as of

2001. This supports stylized statistics presented above suggesting that more than 60% of Romanian migrants eventually returned. The data in Table 4 suggests both migrants and returnees are positively selected over the education variable, relative to the total population. The share of returnees is smallest in the group of people with no degree (and for migrants among those with primary education) while it is largest among those with tertiary education (similarly for migrants). The selection of migrants seems even more skewed towards highly educated relative to returnees. However, these aggregate data hide the already mentioned considerable variation in the selection patterns across destination countries. Nevertheless, even these problematic aggregate statistics reinforce the described migration pattern. The main countries of residence for Romanian migrants according to these data are listed in Table 5. Since it reports the stock of migrants over a longer period of time, the list includes the North American destinations (US and Canada) along with the more recent destinations like Germany, Italy, Spain and Austria.

The main goal of our paper is to analyze the rationality of decisions to migrate and return with regard to the observed selectivity patterns and the sorting of Romanian migrants across these destinations. Using unique sources of data we are able to identify Romanian migrants in three of the main destination countries: the US, Spain and Austria. These countries span very well the type of destination countries for Romanian migrants, each of them corresponding to one of the three selection patterns described above. The US belongs to the group of countries receiving positively selected migrant from Romania, being an Anglo-Saxon country with high returns to skills and relatively unregulated labor markets. Austria is a continental-European country (similar to Germany) with more regulation in the labor markets, lower skill compensation, restrictive immigration policies and it receives, on average, neutrally selected Romanian migrants. Spain (along with Italy) received a large number of negatively selected Romanian migrants and has a rather regulated labor market, with a high concentration of unskilled jobs in construction and manufacturing and thus low skill premia. Our analysis based on individual data will characterize the features of selection for both migrants and returnees. It will relate these to skill-specific premia in order to test if economic rationality is consistent with the observed selection and sorting of Romanian migrants across destinations.

3 Data and Methodology

Following the literature on selection of migrants (e.g. Hanson and Chiquiar 2003, Fernandez-Huertas Moraga 2008) we first characterize the distribution of non-migrants, migrants to different countries and returnees based on their combination of observable characteristics. We group individuals into cells to estimate their wage-earning ability and their probability of employment (in Romania). We call the wage-earning ability, the skill of that group of workers. For each cell, we count non-migrants, returnees and migrants to US, Austria and Spain to determine how these groups compare to each other in their distribution across skills. We define the selection of migrants (positive or negative) as the difference in average skills between migrants and non-migrants. We assess

then if the likelihood of selecting oneself into a group (non-migrants, migrants or returnees) is systematically related to skills.

Our data include wages by each skill cell, both for Romanian migrants (in the US, Austria and Spain) and for non-migrants and returnees (in Romania). We can therefore calculate the average and the skill-specific premium to migrate and to return. Using a simple regression analysis (by skill), we can relate the probability (frequency) of migration/return to the corresponding skill specific premium. Controlling for the costs of migration (that may differ by skill) allows us to test the economic rationality of migration and return. This is a simple modification of the Roy (1951) model to measure selection in many skill groups and to estimate the migration premia with different selection rules, for returnees and permanent migrants.

We describe the individual data and their skill structure in section 3.1. Section 3.2 discusses in detail the measures of average selection on observables. The construction of the average and skill-specific migration and return premium is described in section 3.3. Section 3.4 presents then the model we use in our econometric analysis of the determinants of selection.

3.1 Individual Data and Wage Decomposition

We match information from census data (for employment) and population surveys (for wages) to analyze the characteristics of three groups of Romanian workers around the year 2003: non-migrants, migrants and returnees.

The data for Romania are from the National Demographic Survey (NDS 2003), as well as from the Census 2002. The NDS data were collected by the Center for Regional and Urban Sociology (CURS) and were designed to be representative both at national and regional level. Our restricted sample has more than 35,000 observations, including 1,400 returnees (who had spells of employment abroad), and covers all relevant individual characteristics besides information on migration choices[4]. We use census and income surveys for the three destination countries. For the US, we construct employment, population and average monthly wage data on Romanian migrants by observable characteristics using the 2000 Census. For Spain, we use the 2002 Census for employment and population data on Romanian immigrants and the EU-SILC (2004) for average monthly wage data. For Austria, we use the 2001 Census for employment and population data on Romanian immigrants and the EU-SILC (2004) for the average monthly wage data. We convert all wages into 2003 US $ and we consider that database as a cross section of Romanian individuals circa 2003, either resident in Romania (non movers or returnees) or resident in USA, Austria or Spain. We restrict our sample to individuals between 15 and 65 years of age.

In the constructed data set we observe for each individual i a vector of characteristics X_i and his migration status, i.e. non-migrant in Romania (NM), migrant residing in a destination country c (Mc), or returnee (R) in Romania after an employment spell abroad. Following Chiquiar and Hanson (2005), the vector X includes four

[4] The dataset is described in greater detail in Epstein and Radu (2007).

relevant characteristics defined by the following categorical variables: education (Edu), with the categories No Degree, Primary, Secondary and Tertiary; age (Age), taking ten values from 15 to 65 in 5 years intervals; gender (Gen), with the two categories M and F; and family-size (Fam), with four categories: Single with no children, Married with no children, Single with Children and married with Children. These characteristics identify the observable features of an individual in our dataset. We use the notation $x_i = (Edu_i, Age_i, Gen_i, Fam_i) \in X$ to denote the vector of characteristics of individual i. We allow for the fully saturated model in observable characteristics, so individuals can be put in one of 320 cells spanned by x_i (= 4 education by 10 age by 2 gender by 4 family groups). Each individual has also a "migration status" k_i attached to herself as she can be a non-migrant in Romania, a migrant residing in country c (USA, Spain or Austria) or a returnee, hence $k_i \in \{NM, M_{US}, M_{AUT}, M_{SPA}, R\}$.Our dataset also allows us to actually observe (for Romania and USA) or to impute (for Spain and Austria) based on their occupation and industry, the wage of each individual w_i[5].

We decompose the log wage of individual i working in country j into four components as follows:

$$\ln(w_{ij}) = \ln w(x_i) + \ln p_j(x_i) + I(k_j = R) * \ln r_j(x_i) + \varepsilon_{ij} \tag{1}$$

The term $\ln w(x_i)$ is the mapping from individual observable characteristics x_i into log wages in Romania in 2003. Assuming that the observable characteristics x_i are the main determinants of wage-earning abilities of individuals, the function $\ln w(x_i)$ translates the characteristics into a wage earning potential in Romania. The term $\ln p_j(x_i)$ is the migration premium (or "location" premium as defined by Clemens, Montenegro and Pritchett 2008). It represents the extra wage (in log points) obtained by individual i from working in country j as migrant. The reference country, Romania, will be identified as $j = 0$ and we set, by definition, $\ln p_0(x_i) = 0$. We allow this premium to vary with individual characteristics across skill groups. The term $\ln r_j(x_i)$ is the "return" premium. It is the premium (positive or negative) from being a returnee ($k_j = R$) relative to being a non-migrants I. Finally, ε_{ij} are the idiosyncratic effects on the earning abilities of individual i in country j, which we first assume to have zero-mean in each cell x_i of the set X and to be uncorrelated with x_i, $E(\varepsilon_{ij}|x_i) = 0$. Unobservable wage-earning characteristics of individuals within an observable skill-cell x are thus assumed to be independent and identically distributed with zero average. However, later we will discuss the possibility of non-random unobservable characteristics and its implications on selection issues.

[5] As we do not observe individual wages in the Spanish and Austrian census (and the EI-SILC is too small to have representative wages for Romanian migrants in Austria and Spain), we attribute the average wage based on occupation-industry (from the respective population surveys). The basic idea is that observable characteristics affect the type of occupation-industry in which a person works and the wage is determined by those attributes. In the rest of the paper we will call individual wages the wages constructed following this procedure for Austria and Spain residents. For residents of Romania and US we have the actual individual wages.

3.2 Measures of selection

Our goal is to define and measure two sets of concepts that are crucial to characterize the process of temporary or permanent migration and to elucidate the connections between these concepts. The first set of concepts are the selection of migrants (relative to non-migrants) and the selection of returnees (relative to non-migrants) along the wage-earning ability (skill) dimension. Are migrants (and returnees) selected, on average, among individuals with higher earning abilities (positive selection) or lower earning abilities (negative selection) than the average non-migrants? Given the structure of our data we will be able to characterize the selection of migrants only along the observable wage-earning abilities. This is necessary due to the absence of good instruments but also follows the advice given in Hartog and Winkelmann (2003) against correcting for selectivity when the sample of migrants is small relative to the sample of non-migrants. We will however discuss, in light of the existing literature, what may be the selection of migrants along unobservable skills and how it may affect our findings. As for returnees, we will need an identifying assumption to distinguish selection on unobservables from the return premium. The second concept to be measured is the "premium" from making a migration decision; in particular the premium for being a "migrant" and for being a "returnee". In any economically motivated theory of migration, for a given set of observable characteristics, migrants to a richer country should earn more than non-migrants. However, how does this premium vary with skills and country of destination? Even more interesting would be to know if, for given observable skills, returnees earn more or less than non-migrants. If there is a premium for returnees, then temporary migration has a permanent positive effect on earning abilities. Hence migration and return can be part of a strategy to increase the living standards and returnees are not, on average, those who made a mistake in the initial decision to migrate. Like for the migration premium, it is also very relevant to understand whether and how the return premium depends on skills.

Let us define, in turn the formulas to obtain each of these terms: the average selection of migrants and returnees on observables and the average premium for migrants and returnees, as well as their dependence on observable skill.

3.2.1 Average Selection

The average (logarithmic) wage-earning ability of a non-migrant (NM) with observable characteristics x, call it $\ln \widehat{w}(x)$, is summarized by the average individual wage of all non-migrant individuals in observable cell x. Hence $\ln \widehat{w}(x) = (1/NM_x)\sum_{i \in x} \ln w_{i,NM}$ where NM_x is the observed total employment in cell x. The variable $\ln \widehat{w}(x)$ can be called (wage-earning) skill of group x. The average observed skill of the non-migrants in Romania ("country 0"), corresponds therefore to their average log wage based on observables:

$$\ln w_{NM,0} = \sum_{x \in X} \ln \widehat{w}(x) f_{NM}(x) \tag{2}$$

The term $f_{NM}(x) = NM_x / \sum_{z \in X} NM_z$ is the observed relative frequency of non-migrant workers, NM in cell x. If, conditional on x, the idiosyncratic wage residuals in 1 converge in probability to 0, $(1/NM_x)\sum_{i \in x} \varepsilon_{io} \xrightarrow{p} 0$, then with a large enough sample, such as the census, the value $\ln \widehat{w}(x)$ calculated from the sample would converge to $\ln w(x_i)$. In order to identify how migrants compare to non-migrants in their observable skills, we construct the counter-factual wage distribution based on the observable characteristics of migrants and the corresponding observed wage of non-migrants for each cell x. In particular we define the average skills of migrants to country c, based on observables, as:

$$\ln w_{Mc,0} = \sum_{x \in X} \ln \widehat{w}(x) f_{Mc}(x) \tag{3}$$

The term $f_{Mc}(x) = Mc_x / \sum_{z \in X} Mc_z$ is the relative frequency of migrants to country c, Mc, observed from the census of country c. This method accounts in a fully non-parametric way for the fact that migrants are non-randomly selected from the original population and uses the relative frequencies of migrants to non-migrants to correct for this. Moreover, the differences in skills between migrants and non-migrants are evaluated at home wages, assigning each skill its Romanian price.

Similarly, to identify how returnees to Romania compare to non-migrants we construct the average wage-earning ability of returnees, based on the observable characteristics of returnees and the log wage of non-migrants $\ln \widehat{w}(x)$:

$$\ln w_{R,0} = \sum_{x \in X} \ln \widehat{w}(x) f_{R}(x) \tag{4}$$

Analogous to (3) the term $f_R(x) = R_x / \sum_{z \in X} R_z$ is the relative frequency of returnees in skill cell x. Given the definitions provided above, we define the average "selection"(S) based on observables (O) of migrants to country c, relative to non-migrants as:

$$OS_{Mc,NM} = \ln w_{Mc,0} - \ln w_{NM,0} \tag{5}$$

If expression 5 is positive, migrants to country c are selected on average above the mean of wage-earning characteristics of non-migrants. This is positive selection. Vice-versa, if it is negative, migrants to country c are selected, on average, below the average wage-earning ability of non-migrants. They are negatively selected. Moreover, as the expression is in log differences, it approximates the difference in wage earning abilities as percentage of the average non-migrant wage. Similarly, we define the selection of returnees (on observables) relative to non-migrants as:

$$OS_{R,NM} = \ln w_{R,0} - \ln w_{NM,0} \tag{6}$$

Like above, a value of $OS_{R,NM} > 0$ implies a positive selection of returnees relative to people who did not migrate.

There are two issues that may bias the selection of migrants and returnees according to observable characteristics, produced by 5-6. Those biases may produce the appearance of positive or negative selection when there is none or vice versa. The first issue is that for given observable characteristics participation rates into employment in Romania may be systematically different from participation in the labor market of country c. The second is that there may be unobserved characteristics correlated with the x (hence not random and not zero-mean within group x) and those may differ between migrants and non-migrants. We will discuss them in turn.

3.2.2 Participation into employment and unobservable characteristics

The rate of participation into employment for a group with characteristics x can be different at home and abroad. It is easy to think that if a skill group x is paid a higher wage in a country this may attract workers of that skill group to that country and push a larger fraction of them to work. This may affect the calculated skill selection, if we base our evaluation of formulas 5 to 6 on employment data. For instance, if migrants to country c have characteristics that are identical to non-migrants but, once in the labor market of country c, their participation to employment is relatively larger in the high wage-potential groups compared to their participation in Romania, the method above will produce the appearance of positive selection, when there is in fact no selection. Had those migrants stayed in Romania, they would have earned, on average, as much as non-migrants. To avoid this problem, we should correct the relative frequency of migrants in constructing their average wage earning ability $\ln w_{Mc,0}$. In particular, rather than the frequency of characteristic x in employment we should use its frequency in the *population* of migrants and correct those population frequencies by the participation rates of each group x in Romania. Such correction allows us to compare the average wage-earning ability of migrants, had they stayed in Romania, with that of non movers. Formally we can define the "participation-corrected" average wage earning ability of migrants to country c as follows:

$$\ln w_{Mc,0}^{PART_0} = \sum_{x \in X} \ln \widehat{w}(x) f_{Mc}^{PART_O}(x) \tag{7}$$

where $f_{Mc}^{PART_O}(x) = \theta_x^0 Mc_x^{POP} / \sum_{z \in X} \theta_z^0 Mc_z^{POP}$ and $Mc_x^{POP}$is the total population (rather than workers only) with characteristic x who migrated to country c, while θ_x^0 is the employment-population ratio for workers of characteristic x in Romania ($\theta_x^0 = NM_x/NM_x^{POP}$). We will use the empirical participation rate of non-migrants in each cell from the Romanian Census 2002, as non parametric estimate of θ_x^0, and the data on the population Mc_x^{POP} of migrants in group x in country c from the Census of country c. The "double selection" into migration and into employment that is considered in many recent papers (e.g. Chiquiar and Hanson 2005, Fernandez-Huerta Moraga 2008, Piracha and Vadean 2009) is addressed here in a completely non-parametric way. Assuming that we have identified the relevant observable characteristics that determine the probability of migrating and of participating into the labor force, we use a fully non-parametric relation between those and the

migration probability, and between those and participation at home, to identify the selection on wage-earning abilities. In particular the variable:

$$OS^{PART_0}_{Mc,NM} = \ln w^{PART_0}_{Mc,0} - \ln w_{NM,0} \tag{8}$$

identifies the difference in wage-earning ability of migrants had they remained at home relative to the wage-earning abilities of non-migrants. This is the cleanest comparison possible to identify the type of migrant selection on observable wage-earning abilities. Similarly, we can correct the skill selection of returnees by imputing to them the employment-population ratio of non-migrants.

3.2.3 Unobservable characteristics

The unobservable individual characteristics denoted as ε_{ij} in expression 1 have been assumed to be uncorrelated with x so that $E(\varepsilon_{ij}|x) = 0$. However, it is possible that some unobservable characteristics are correlated with x so that $E(\varepsilon_{ij}|x) = g(x)$. For instance if unobserved wage-earning abilities are larger, on average, for groups with larger observable wage earning ability then $g(x)$ can be systematically positively correlated with $\ln w(x)$. Under these circumstances the term $(1/N_x)\sum_{i\in x}\varepsilon_{io}$does not converge in probability to 0 and hence cannot be approximated to 0 using the Census sample. In fact, if different selection processes operate on the unobservable characteristics it may even be possible that: $E(\varepsilon^{M_c}_{io}|x) = g^{Mc}(x) \neq E(\varepsilon^{NM}_{io}|x) = g^{NM}(x)$ which means the conditional average of unobservable wage earning ability for a group x is different between migrants and non-migrants.

This departure from the original assumptions implies that the total average skill selection indicator $S_{Mc,NM}$ will equal:

$$\begin{aligned} S_{Mc,NM} = OS_{Mc,NM} + US_{Mc,NM} = \\ \ln w_{Mc,0} - \ln w_{NM,0} + \sum_{x\in X} g^{NM}(x) f_{NM}(x) - \sum_{x\in X} g^{Mc}(x) f_{Mc}(x) \end{aligned} \tag{9}$$

Where the term $OS_{Mc,NM}$ is constructed as in expression 5 and is the selection based on the observables, while the term $US_{Mc,NM} = \sum_{x\in X} g^{NM}(x) f_{NM}(x) - \sum_{x\in X} g^{Mc}(x) f_{Mc}(x)$ is capturing the selection of migrants over the unobserved wage earning abilities. The term $US_{Mc,NM}$ cannot be constructed with our data. To do this one would need information on the wage paid to migrants in Romania, before they migrated. Some recent studies on Mexican data (Fernandez Huertas-Moraga 2008, Kaestner and Malamud 2010) have used data on pre-migration wages and have evaluated such a term for Mexican migrants. Clemens et al (2008) also evaluate this for the Philippines, South Africa and Mexico. These are countries not too far from the income level of Romania, hence we can look at the average selection of migrants on unobservable skills there, especially relative to selection on observables, to gather an idea of how large that phenomenon could be. While it is hard to have

a clear theoretical expectation on the sign and magnitude of the selection on unobservables, two considerations may help.

Why should migration costs or migration selection by the receiving country be strongly related to some unobserved abilities? First, while in some specific cases certain skills might affect migration behavior (e.g. knowing one specific language), it is hard to see how these are systematically correlated with observables and, in the aggregate population, this may not matter much. Second, if we consider an economic rationale for migrating, the type of selection produced on observables should be the same (positive or negative) as on unobservables. A country that rewards wage-earning skills would attract more skilled workers along the observable and unobservable dimension. In accordance with this intuition most of the existing estimates of observable and unobservable selection either find no relevant selection on unobservables (Kaestner and Malamud 2010) or find selection on unobservable of the same sign and smaller scale than selection on observables (Budnik 2009, Fernandez Huertas-Moraga 2008 and the relevant cases in Clemens et al 2008).

3.3 Income premia and selectivity patterns of migration

A similar non parametric method can be used to identify, under some assumptions, the average premia, both for migrants and for returnees. Consider the counter-factual wage (4) that returnees would earn if they were paid as non-migrants, with the same characteristic x and difference between this and their actual average wage. This difference represents exactly the average premium to returnees (call it "$PR_{R,0}$") plus a term representing the selection of migrants on unobservables:

$$\sum_{x\in X} \ln w_R(x) f_R(x) - \sum_{x\in X} \ln w_{NM}(x) f_R(x) = \sum_{x\in X} \ln r(x) f_R(x) + US_{R,NM} = \tag{10}$$
$$= PR_{R,0} + US_{R,NM}$$

The term $\ln r(x)$ (from the decomposition of individual wages in expression 1) is the "return" premium for being a returnee and may depend on x. On the other hand, if returnees differ systematically on unobservables from non-migrants then there would be an extra term $US_{R,NM}$ capturing the selection on unobservables. Assuming unobservable wage earning skills of returnees relative to natives are independent of x, then $US_{R,NM} = 0$ so that the expression above defines $PR_{R,0}$.

Finally, we can compute the wage premium that the average migrant to country c will receive relative to what she would have earned at home. This is the "migration" or "location" premium, i.e. the fact that the receiving country pays more for given observable characteristics than the home country Romania. The average premium to migrate to country c (plus the selection on unobserved characteristics) is calculated using the observable

characteristics of migrants to that country as:

$$\sum_{x\in X} \ln w_{Mc}(x) f_{Mc}(x) - \sum_{x\in X} \ln w_{NM,0}(x) f_{Mc}(x) = \sum_{x\in X} \left[\ln p_c(x)\right] f_{Mc}(x) + US_{Mc,NM} = \tag{11}$$

$$PR_{M,c} + US_{Mc,NM}$$

Notice that the term $\ln w_{Mc}(x)$ is the wage earned in country c by Romanian immigrants of skill x. Using the individual wage definition in 1, the difference in the wage of an individual with characteristic x earned at home, 0, or abroad, c, is the sum of the individual location premium, $\ln p_c(x)$, weighted by the frequency of Romanian migrants to country c plus the unobserved selection of migrants to country c, $US_{Mc,NM}$. Given the lack of information on $US_{Mc,NM}$ we will consider it as relatively small, vis-a-vis $PR_{M,c}$ so that we can neglect it and the expression 11 identifies the average migration premium.

3.4 Skill Premium and Skill-Selection

Section 3.3 defines some aggregate statistics to characterize the selection and the premium for migrants and returnees. Being based on the partition of the population into cells $x \in X$, this method defines also the selection and the premium for each value x. Even more conveniently, as the function $\ln \widehat{w}(x)$ transforms the multidimensional set of characteristics X into a unidimensional skill, $\ln w$, we can invert the mapping $(x^{-1}(\ln w))$ and define selection and premia for each level of the skill variable $\ln w$[6]. In particular, using the notation introduced in section 3.3 the selection of migrants relative to non-migrants is measured as a function of the wage level by the relative density: $f_{Mc}(x^{-1}(\ln w))/f_{NM}(x^{-1}(\ln w))$. For instance, a value of this relative frequency for a cell equal to 1.3 implies that in this cell people are 30% more likely to migrate relative to staying, than in the average cell. A value of 1 implies that in the cell people have the average probability of migrating to c. Similarly, the selection of returnees relative to non-migrant over the skill spectrum $\ln w$ is given by: $f_R(x^{-1}(\ln w))/f_{NM}(x^{-1}(\ln w))$. The logarithmic premium for migrants at each level of skill can be written as: $PR_{Mc}(x^{-1}(\ln w)) = \ln w_{Mc}(x) - \ln w_{NM}(x)$ and similarly $PR_{R0}(x^{-1}(\ln w)) = \ln w_R(x) - \ln w_{NM}(x)$ where the wage differences are taken for workers of same skill x.

The representation of selection (relative frequency) as a function of skills is helpful to illustrate the whole profile (kernel distribution) of each group (non-migrants, migrants and returnees). Similarly, the characterization of the premia as a function of skills, $\ln w$, allows us to analyze more systematically how they are related.

In a very simple theory of migration, however, it is also useful to consider each skill cell $x \in X$ as an observation on a group of workers (whose number is equal to population in the cell) who have specific characteristics. Assuming each group as having a random distribution of migration costs to each country and a common return from migration to country c which is given by the common linear premium $LPR_{Mc}(x) = w_{Mc}(x) - w_{NM}(x)$

[6] Such an inversion require the the mapping to be one to one. While, in theory several cells could map to a single wage, in practice as the wage valuse are continuous each cell maps only to one wage value satisfying the requirement.

under general assumption on the distribution of costs the odds of migrating to that country relative to non migrating are an increasing function of the linear premium. Allowing for a measurement error $u(x)$ in the relative frequencies, this can be approximated by the following linear relation:

$$f_{Mc}(x)/f_{NM}(x) = a(x) + b * LPR_{Mc}(x) + u(x) \text{ for } \ x \in X \tag{12}$$

The relative selection in group x indicates by how much the migrants are over (>1) or under (<1) represented in that skill group relative to non-migrants. Two qualifications are needed. First, under the assumption of idiosyncratic costs distributed as an extreme value Gumbull distribution, the standard utility maximization in the Logit model implies that there is a linear relation between log odds and wage differentials (see for instance Ortega and Peri, 2009). Expression 12 is simply a linear approximation of that exact equation. Second, the coefficient b captures whether the selection, consistent with utility maximization, is increasing in the linear returns to migration. The term $a(x)$ introduces the possibility that the selection is affected also by migration costs that are systematically different by skill group. Regression 12 will be estimated for each country of emigration to see if the implication that $b > 0$, derived from a model of rational migration, is supported in the data. In testing the equation for each country of emigration we are assuming *independence from irrelevant alternatives*. Similarly, as we have an independent measure of return premium, $LPR_R(x) = w_R(x) - w_{NM}(x)$ for each skill group, we can test whether the data support a theory of return motivated by economic benefits. We will run the regression:

$$f_R(x)/f_{NM}(x) = \alpha(x) + \beta * LPR_R(x) + v(x) \text{ for } \ x \in X \tag{13}$$

and test for $\beta > 0$. People need not return to a wage equal to that of similar non-migrants. In this perspective migration and return can be the optimal choice, even with no uncertainty (or unexpected shocks) for some people, as we will see in section 6.

4 Evidence on Selection and Premia

To analyze the evidence on selection and premia for Romanian individuals in year 2003,we first show some simple graphs of selection for migrants and returnees over education and age. We will then present the values of the average skill selection on observables as well as the whole distribution of skills for migrants and returnees relative to non-migrants. Finally, we will show the average migration and return premium and their distribution by skill for migrants and returnees

4.1 Simple selection on Education and Age

Figure 1 and 2 present aggregate evidence on the selection of returnees and migrants to each of the three destination countries over education and age groups. Each panel of Figure 1 shows the distribution of non-migrants and one other group (in turn returnees and migrants) in the form of histograms over four education groups (no degree, primary, secondary and tertiary). The wider bars represent the distribution of non-migrants, always the comparison group, and the thinner bars the distribution of the other group. Figure 2 does the same for the distribution across age groups. Panel 1 reports the comparison with returnees, Panel 2 with migrants to the US, Panel 3 with migrants to Austria and Panel 4 with migrants to Spain. In each panel the distribution, which is relative to working individuals (male and female), has been constructed using census data. Some tendencies are already clear from these figures and anticipate some of the regularities that we will unveil later. First, returnees are clearly positively selected among education groups vis-a-vis non-migrants. Their relative distribution is much more skewed towards workers with tertiary education at the expenses of workers in any other education group. In terms of age, returnees are much less differentiated from non-migrants, however tend to be slightly over-represented among groups with intermediate and old age rather than among young workers (below 25). Migrants to the US tend to be better educated as well as older relative to non-movers. Both features may add to their earning abilities. The largest share of migrants to the US is among workers with secondary schooling and above, and they are significantly over-represented among workers older than 50. Migrants to Austria seem the group with the more "average" selection relative to non movers. Their education distribution is not very different from that of non-movers (except for a slightly larger share of secondary educated and a smaller share of those with no degree). The age distribution is only slightly more concentrated in the group 30 to 50 relative to non-migrants. Finally, migrants to Spain show a clear "negative" selection, being much more concentrated than non-migrants among workers with only a primary degree (across education groups) and in the groups of less than 30 years of age (among age groups).

To summarize, the observable feature of returnees look similar to that of migrants to the US, the group of returnees showing the strongest educational distribution. Migrants to Austria, on the other hand, are the most similar to non movers and they show a concentration in intermediate education and age groups. Finally, migrants to Spain seem the group with lowest earning potential skills as they are concentrated among low education and young age groups. We will test more formally in the next section whether these stylized facts match the more carefully constructed measures of average selection.

4.2 Selection on observable wage-earning skills

Table 6 shows the values of the average skill selection of returnees and migrants to the US, Austria and Spain relative to non-migrants. The entries in Column (1) of Table 6 are (respectively from the first to the last row) the statistics $OS_{R,NM}$, $OS_{MUS,NM}$, $OS_{MAut,NM}$, $OS_{MSpa,NM}$ defined as in section 3.2. In column (1) we

construct the frequencies for the group of non-migrants $f_{NM}(x)$ using the Census 2002 data. In column (2) we evaluate the same statistics when the frequencies $f_{NM}(x)$ are measured using the NDS 2003. Column (3) shows the average selection statistics obtained when we correct for participation in the destination country using the observed participation in Romania. Column (4) shows the statistics obtained using only employment and wages of male workers. Column (5) removes from the Romanian sample the ethnic minorities (Roma) who may be significantly different in their wage earning ability from the ethnic Romanian. The values can be interpreted as percentage differences in the average wage earning skill of the group and the average wage-earning skills of non-migrants.

The statistics obtained using different methods and samples show only small variations. This reinforces the idea that the features of selection that we found are quite robust and stable. First, the group of returnees exhibits a positive average selection between 12 and 14%. This means that when compared to non-movers, returnees have observable skills that allows them to earn domestic (monthly) wages higher by 12-14%. This is a significant positive selection. To give some point of comparison, the Mincerian returns to schooling that we estimated on the Romanian NDS data give a return around 0.06-0.07 per year of schooling. Hence, the average difference in skills between non-migrants and returnees is equivalent to 2 years of schooling. This value is not very sensitive to the corrections. Importantly, the number obtained when using the NDS employment data and the number obtained when using employment from the Census are very similar, implying that as far as the selection of returnees is concerned the two data produce compatible results.

Moving to the average selection of migrants to the US we also find a large and economically significant positive selection ranging between 0.13 and 0.20. The only correction that makes some difference is the one for participation which actually increases the selection, implying that the selection of individuals who migrate to the US is even more positive that the selection of working individuals. This may be due to a lower participation of more educated women to employment in the US if they move e.g. with their highly educated working husband. Again, the pure skill selection among these migrants makes them equivalent to workers with 2-3 more years of schooling than the average non-migrant. Confirming the first impression from the education and age data, the selection of migrants to Austria is essentially zero. The statistic is small implying at most a 2-3% positive selection. Migrants to Austria are selected in a way that is not much correlated with their wage-earning skills. Correcting for participation in Romania and using of the NDS rather than the Census data to construct employment frequencies does not make much difference. Finally, the migrants to Spain exhibit a significant negative selection. Confirming the evidence from the education and age data, their average skill selection ranges from -0.07 to -0.13. Using participation rates in Romania (column 3) reduces slightly the negative selection, which implies that Romanian migrants to Spain also have lower employment participation in higher skill groups. Migrants to Spain have skills equivalent to one to two fewer years of schooling relative to Romanian non-migrants.

The average values of the selection variable conceal a whole distribution of skills for each group relative to

non-migrants. Figures 3 and 4 show the comparison for the whole density distribution of non-migrants and other groups. Figure 3 shows the comparison between non-migrants and returnees. We show the distribution of the two groups by skill (logarithmic monthly wages). Two differences are clear even to a cursory visual inspection. First, the density of returnees is consistently lower in the skill range corresponding to 400$ to 1000$ (monthly). On the other hand, the density of returnees is larger for wages above 1000$ and has a particular peak of density around 1600$. These workers are likely to be the college educated in some intermediate age groups. Overall we can reject the hypothesis that the two distribution are equal by doing a Kolmogorov-Smirnov test, which rejects equality at 0.1% significance. Figure 4 shows the kernel density estimator for non-migrants and migrant in each of the 3 destinations both using employment distribution by skill (Panel 1) and population distribution (Panel 2). The solid line represents non-migrants, the short dashed line is for migrants to Austria, the long dashed line for migrants to Spain and the dotted line for migrants to the US. As expected, relative to the non-migrants, the distribution of migrants to Spain shows a significant density mass below the average skill level of non-migrants (about 882 $) with a peak near 700 $. On the other hand the distribution of migrants to the US shows a significant mass of density above the average of non-migrants reaching high and very high wages (up to 1800 $). The density of migrants to Austria is not too different from that of non-migrants. A Kolmogorov-Smirnov test of distributional equality between non-movers and migrants to Austria cannot reject the null at 5% confidence.

All in all, average skill selection on observables ranges from -13% for migrants to Spain to +16% for migrants to the US averaging around 0 for migrants to Austria. It is quite hard to say how much and in what direction the unobserved selection would modify these numbers. In comparison Huertas-Moraga (2008) who estimates negative selection for migrants from Mexico to the US reports that selection on unobservables is also negative and about 30% of selection on observables. Kastner and Malamud (2010) do not find any significant selection either on observables or on unobservables for the same Mexican migrants to the US. Clemens et al. (2008) report a selection on unobservables for migrants from the Philippines equal to 8% and for South Africa they report and even more positive selection on unobservables (around 20%). The few other estimates available are for much poorer countries. In general, previous studies have either found an average selection on unobservables of the same sign as the selection of observables but much smaller or no selection at all. With this caveat in mind, we interpret the average observed selection as a correct measure of skill selection and proceed to identify the migration and return premium.

4.3 Migration and Return Premium

The largest economic benefit of international migration is in form of a "migration premium" for migrants. Individuals with given skill characteristics increase substantially their wage and income by moving to countries where their skills are paid much more. The average wage premium for migrants varies across countries of destination, but so does the skill-profile of migrants, which depends on how the labor market at the destination price their skills. In general, for a given average wage differential, the influential Roy (1951) model (applied for

instance in Borjas 1987 and Borjas and Bratsberg 1996) implies that countries with large skill compensation (namely larger than in the country of origin) attract more skilled workers. Those countries typically exhibit larger wage inequality driven by skill differences. To the contrary, given average wage differentials, countries with low skill compensation (lower than in the country of origin) would attract instead less skilled workers. In the first case, this positive selection results from the migration premium increasing in skills while, in the second case, negative selection results from the migration premium decreasing in skills.

A simple way of characterizing such migration premia across skills is to report the distribution of the log wages earned by migrants abroad and those wages migrants would have received at home (imputed based on their observable characteristics). Averaging those two distribution using the density of skills of migrants and taking their difference would generate the average migration premium. The distributions of wages in the country of emigration together with what those individuals would earn in Romania is shown in Figure 5. The difference in the average skills between the two distributions represents the average migration premium and is reported in 2003 US $ below each panel. Panel 1 reports the wage distribution for migrants in the US and their counterfactual distribution had they worked in Romania. Panel 2 shows the same comparison for migrants to Austria and Panel 3 for migrants to Spain. Two regularities are apparent. First, relative to their wage distribution in Romania, migrants have a wider wage dispersion in the US, an intermediate one in Austria and the smallest one in Spain, smaller even than in Romania. This is an indication that returns to skills are highest in the US and lowest in Spain. Second, while significant in each case, the average migration premium is much more substantial for migrants to the US (990 $ per month) than for migrants to Spain (300 $ per month). This is consistent with the large migration flows to the US, and it partly compensates the large costs of migrating to that country (farther and its more different culture). More interestingly, however, is the fact that for migrants to Spain the figure suggests that the largest benefits would accrue to those who are in the long left tail of the counter-factual Romanian wage distribution (hence the low skilled). To the contrary, for the migrants to the US, the more likely to gain are those in the right tail of the wage distribution. A more systematic analysis of premium and skills is needed here, although the simple wage distribution already indicates the main driver of migration incentives between these countries.

5 Migration and return driven by skill-specific premia

In this section we analyze the migration and return premium in relation to skills and estimate the regressions from section 3.4 to identify the correlation between migration and premia and to test the rationality of migration. The upper part of Table 7 shows the correlation between premium and skills for, respectively, returnees (Column 1), migrants to the US (Column 2), to Austria (Column 3) and to Spain (Column 4). The first three rows show the correlation between the linear premium and the skills ($\ln w$) across skill groups (the units of observation are the 320 x cells). The regressions are estimated using least squares weighted by employment in the cell, in the

first row. In the second row we control for age and in the third row we add family-type fixed effects to check whether within an age-family type group the premium for migration and return changes with skills (mainly education) and whether the sign is as in the overall regression. The results are quite robust and confirm the simple evidence from the previous section. The premium for returnees and migrants to the US is very strongly positively correlated with skills. The premium for migrants to Austria is neutral in the skill dimension; namely it has no systematic correlation with the skill level of a group. There are certainly some groups for which the migration premium is higher than others, but this difference is not systematically related to $\ln w$. Finally, the premium for migration to Spain is negatively related to skills: lower skills receive a larger wage-premium for migration. As the dependent variables are in thousands of 2003 US $ and the explanatory variable is in $\ln w$, we can interpret the coefficient as a semi elasticity. For instance, an increase in skills by 10% (equivalent to about 1.6 extra years of schooling, given a Mincerian return of 0.06 in Romania) would imply an increase in return premium by 36 to 49 US $ for returnees, an increase by 114-128 $ of the premium for migrating to the US, no change in the premium for migrating to Austria and a decrease of 44 to 75 $ in the premium for migrating to Spain. Such effects are significantly different from each other and precisely estimated.

The lower part of Table 7 characterizes the linear premia in relation to education levels. We regress the premia on three dummies for Primary, Secondary and Tertiary schooling. The omitted dummy is "No degree" and the estimated coefficients are reported in rows 4 to 6. Interestingly, for the return premium, the largest estimated dummy is for college educated, while the premium for primary and secondary educated is not very large (relative to 0, the premium for those with no degree). Hence, simply isolating the education dimension, most of the positive correlation of the return premium with skills derives from college educated. The case of the premium for migrating to the US is different, since all three education groups receive a significant premium relative to the group with no education. The premium for college educated is only marginally larger than the premium for primary educated. For migration to Austria there seems to be a negative premium to primary educated but a positive one for college educated. This non-monotonic effect may be part of the reason that we do not estimate a clear dependence of the premium on skills for migration to Austria. Finally for migration to Spain the largest premium of all is for primary educated, and it is much lower for secondary and tertiary educated, explaining the negative relation.

Table 8 shows the estimates of coefficient b and β from equations 12 and 13 in its first three rows. Do larger wage premia bring higher frequencies of migrants and returnees? If the estimated coefficient is significantly positive, no matter what is the structure of the premium over skills, then migrants and returnees respond to that premium by skill group proving the rationality of their behavior. The estimates clearly show that within each skill group, the higher the premium to return and migration, the much higher will be also the relative frequency of return and migration. In the third row we also control for a full set of age and family structure dummies. These dummies are meant to capture differential migration and return costs for individuals of different age groups and different family structure. Young, unmarried individuals with no children are the most mobile,

hence one can expect that in these groups we observe the most migrants and returnees beyond the effects of a wage premium. This would be due to a systematic difference in costs rather than in the return to migration. The inclusion of these proxies for migration costs only affects the estimates of the coefficient for Spain, which turns insignificant. The other cases maintain a positive and significant correlation between returns and migration frequency. The estimated b and β coefficients are always positive and significant in 11 cases out of 12. Their value ranges between 0.1 and 0.6 with most estimates between 0.2 and 0.4. Taking 0.25 as the median estimate, this coefficient implies that an increase in the migration premium for a skill group by 1,000 $ per month would increase the frequency of migrants relative to non-migrants in that skill group by 25%. The stability of the coefficient across countries and even between migrants and returnees implies that we can think of a common explanation for the skill selection of migrants and returnees, namely their response to the wage premium, i.e. to economic incentives. The different skill composition of migration to different countries and the skills of returnees can be explained simply by the common tendency to migrate where and to return when the premium is larger. This common response to incentives is consistent with a positive skill-selection for returnees and migrants to the US, with a negative selection for migrants to Spain, but also with the neutral selection of migrants to Austria. These migrants too respond to wage premia, only those premia do not have a clear correlation with skills. The last three rows of Table 8 report the correlation of return or migration frequencies with education dummies and confirm the described selection pattern.

6 Implications in a Model of Education, Migration and Return

There are two notable results in the previous empirical analysis. First, returnees to Romania are clearly positively selected relative to non-migrants. It is harder to say whether they are positively selected from all migrants but their positive selection is similar to migrants to the US. Second, returnees earn wages significantly higher than non-migrants and this difference increases with their skills. Interpreting the wage premium as a productivity difference due to skills accumulated abroad, there are two potentially important effects of migration and return for the sending country. First, this process may increase the expected overall returns to skills, possibly inducing the positive brain gain incentives emphasized by Docquier and Rapoport (2008) and offsetting a negative brain drain. Second, it may increase the productivity of returnees (who have learned new skills and enhanced their human capital) with positive effects for the domestic economy.

As the evidence points to a rational migration behavior, driven by migration and return premia, we use a simple model (developed in Mayr and Peri 2009) to inquire more systematically the size of these two effects for Romania. Disciplining the model using the estimated parameters for Romania, we quantify the effects of migration in terms of years of schooling and average wages and simulate the effects of relaxing migration constraints. We will briefly describe the intuition of the model below. The details of the model solution and of the parametrization of the model can be found in Mayr and Peri (2009). The intuition of the model is simple

and will guide us to identify the simulated effects of freer migration on schooling and wages once we account for return.

6.1 Key assumptions and Results of the model

We consider the Romanian economy as the home country and indicate it with H. Romanians live two periods. In the first they pursue education and then decide whether to migrate and work. In the second period they return or stay abroad. The wage of a Romanian with schooling h_i at home, in the first period is:

$$\ln(w_{Hi}^1) = \ln(w_H^{NS}) + \eta_H h_i \tag{14}$$

where $\ln(w_H^{NS})$ is the domestic wage of the worker with no schooling (NS). We assume that the agent's utility function is separable over time and it is logarithmic in each period's income so that expression (14) also represents the period utility from working and living at Home. If the individual migrates to a Foreign country (F) his wage is $\ln(w_F^{NS}) + \eta_F h_i$. At the same time we assume that there are costs of living abroad (material as well as psychological) and that those costs are specific to the period of the individual's life. We express these costs in utility units and denote them by M_1 and M_2 where the subscripts refer to the period in which they are incurred. Hence, the utility abroad (logarithmic wage net of costs of living abroad) for individual i when young is:

$$\ln(w_{Fi}^1) - M_1 = \ln(w_F^{NS}) + \eta_F h_i - M_1 \tag{15}$$

If the individual chooses to remain abroad in the second period, she will receive the following utility (logarithmic wage net of costs of living abroad):

$$\ln(w_{Fi}^2) - M_2 = \ln(w_F^{NS}) + \eta_F h_i - M_2 \tag{16}$$

As Romania is poorer than the average country of emigration, $\ln(w_F^{NS}) > \ln(w_H^{NS})$. In the case of migration to a country, such as the US, where the returns to schooling are higher, then $\eta_F > \eta_H$.

Romanians who have been abroad for one period "enhanced" their human capital by learning new skills. If they decide to return, this would increase their earnings per unit of initial human capital (as an augmentation of their human capital). Moreover this premium, according to the evidence in the previous sections, is increasing with skills. Hence, the (logarithmic) wage of a person who returns to the home country in the second period of her life after having been abroad is:

$$\ln(w_{FHi}^2) = \ln(w_H^{NS}) + \eta_H h_i + \kappa h_i \tag{17}$$

where w_{FH}^2 indicates the wage in the second period of life (superscript) for individual j who has been abroad

and returned. The parameter $\kappa > 0$ is the extra return for human capital associated with the experience abroad. Finally, the utility of workers who stayed at home is identical in the first and second period and is given by the following expression: $\ln(w_{Hi}^2) = \ln(w_H^{NS}) + \eta_H h_i$.

The individual decisions are as follows. At the beginning of the first period (youth) individual i chooses how much schooling to get, h_i, and simultaneously pays the cost, k_i, for this education. We assume that this cost of education is positively related to the amount of schooling achieved, h_i and inversely related to some individual skills ν_i, so that $k_i = \frac{\theta h_i^2}{\nu_i}$ and θ is a common cost of getting education. In equilibrium the optimal amount of schooling h_i is a monotonically increasing function of the skill ν_i and schooling perfectly reveals individual skills. Immediately after their schooling decision (still at the beginning of period 1) the individual chooses to migrate or to stay. We treat migration as a lottery and the decision to participate in the lottery as voluntary. Once an individual has entered the lottery she faces the same probability of migrating as any other participant $p \in [0, 1]$. This lottery is our way of capturing migration openness. The probability p has to do with a rationing of migrants from the receiving country point of view. A policy of receiving countries that open the borders to all immigrants would result into $p = 1$. The regime before the collapse of the Soviet Block corresponded essentially to $p = 0$. At the beginning of the second period people who remained at Home continue to earn wage w_{Hi}. We assume that the cost of moving in the second period is too high to make it profitable (or that the receiving country has a policy which significantly penalizes the immigration of older workers), while emigrants living abroad can decide whether to stay in Foreign or to return.

The solution of the model[7] identifies the selection of migrants and returnees, in terms of their schooling h_i (and the underlying skill parameter ν_i).This simple (log linear) structure of wages, utility and costs implies that the model produces some "threshold" skill levels. The key parameter condition is as follows. If $\kappa + \eta_H \geq \eta > \eta_H$ the return to migrating are higher for highly educated, hence migrants are relatively more educated. All workers with skills (schooling level) above a threshold h_M will enter the migration lottery (and only a fraction p of them will actually migrate). However the returns to return migrating are even higher for highly educated and hence the most educated of all choose to migrate and return. In particular there will be a higher schooling threshold h_R above which all individual, if migrated in the first period would return in the second. Hence those with intermediate schooling (between h_M and h_R) choose to migrate and stay abroad (if they succeed to migrate), least educated (below h_M) stay at home. The most educated (above h_R) migrate and return. The model has one important implication. If the probability of migrating, p, increases under positive skill selection more intermediate and high skilled will migrate. However, two effects may balance this brain drain. First, as education is a choice, more individuals will choose higher education as the expected returns to schooling have increased. Having higher probability to migrate (and return) increases the expected returns to education and induces more individuals to get higher education[8]. Second, more migrants means more returnees and each one of them will

[7] For the details of solution see Mayr and Peri (2009).

[8] The response of education depends on the assumed costs of education and distribution of skills that we have set to match the

benefit from the extra-productivity (wage) effect due to the accumulated skills aboard. These two positive effects on skills and wages can in part or completely offset the negative effect of a positive selection of migrants that reduces the skill of those left in Romania.

The migration and return costs are set to match the share of returnees in total (always measured to be around 0.4-0.5). The wages for no schooling $\ln(w_F^{NS}), \ln(w_H^{NS})$ are set at the level observed from our Romania data and the average of the three migration countries. The parameters η_F and η_H are the returns to schooling estimated using a Mincerian equation for Romania (around 0.06) and for the average European country (around 0.08). The parameter $\kappa = 0.025$ is chosen to match the return premium obtained by college educated returnees (around 0.28 over non-migrants). The other parameters of the model are kept as in Mayr and Peri (2009) to match an average Eastern European Country.

6.2 Migration, Return and average wages in Romania

Table 9 shows the simulated effects on years of schooling and wages (standardizing the wage with no migration to 1) when we increase the probability of migration from 0 to 0.30 by increments of 0.05. To match the current percentage of Romanians abroad as of 2003, the value of p would be near 0.10. For that value returnees equal 4.5% of the population in Romania and migrants (still abroad) equal 4.6% of Romanians. These numbers are not too far from the 5% of returnees found in the NDS and the 3.2% of migrants in the Docquier and Marfouk (2006) data. The first result of the simulation, shown in the first three rows of Table 9, is that relative to the case of 0 migration the schooling of young and old people increases as the probability of migration increases. In spite of having a loss of highly educated young workers due to positively selected migration the incentive effect on schooling more than balances this tendency. The effect on average schooling of young individuals is thus purely an incentive effect. The effect on old workers, to the contrary, combines also the positive selection of returnees, so that as the most educated come back this further increases the average education of old relative to the case with no migration. Overall, the effect is that average schooling in the population is higher by half a year due to international mobility relative to the case with no migration. With a probability of migration equal to 0.20, which is double the current value, There would be an increase of average schooling of Romanian population by one year relative to the case with no migration.

The wage effects, reported in rows 4 to 10 are also interesting. Again the effect on young is purely the education incentive effect. It amounts to a plus 2% (when $p = 0.10$) relative to no migration and it could be increased to +5% for $p = 0.20$. The effects on Old workers combines the incentive and the return premium and generates an average increase by 9%, relative to no migration (for $p = 0.10$). That gain increases to +22% by doubling the migration flows (loosening the policy to $p = 0.20$). The rows showing the effect on wages by schooling level (wages are always relative to average wage with no migration) also show that all the gains from

initial distribution of Romanian population by schooling level (from the Barro-Lee 2000 data).

migration and return accrue to the high schooling group which is the one most affected by the positive incentives and most rewarded by the return premium. The less educated are those who do not migrate in the model, hence there are no effects for them. The intermediate education is the group that migrates but does not return, hence the return premium has no incentive effect on them nor has a direct effect in raising wages. The wage of highly educated, however, rises by 44% for young individuals (due to their much larger education) and by 58% for old individuals, due to higher schooling and the return premium.

If we were to eliminate the schooling incentive effect from the simulation (table available upon request) we would observe a negative schooling and wage effect of migration on the young generation (due to brain drain) but still a positive schooling and wage effect on the old generation (brain return and return premium). The two effects in our model would still give a small positive average wage effect.

6.3 Some Policy Experiments

The model also allows us to change, by some small amount, the policy variables and analyze what would happen to migration, return and wages in Romania[9]. For instance a decrease of the variable M_1, the cost of residing abroad, possibly associated with better communication, more homogeneous institutions, more integrated markets implies higher migration, lower percentage of returnees and thus a stronger incentive-effect of migration increasing the schooling level. And just the opposite effects would occur if costs increased. For an increase of these migration costs by 20% average wages and average years of schooling in Romania would decrease by 1%. More interestingly, an increase in the reward to highly educated in the receiving countries, such as a fiscal incentive for high earner in these countries, or a larger decrease in tax rates for highly educated, would generate two opposite effects. On the one hand, as potential emigrants will have larger expected returns to skills in the receiving country, this will produce a stronger education incentive effect in the sending country and will increase average schooling of young. On the other hand, fewer migrants would return, and this would decrease schooling of the old generation in Romania. For a fiscal incentive that only increases returns to schooling from 8% to 8.3% in the receiving country, the two effects will produce a 2% wage and schooling increase in Romania all due to higher education of young. Finally, the country of origin may introduce some fiscal incentive for highly skilled returnees, corresponding to an increase in the parameter κ. A fiscal incentive that increase by 30% the premium of returnees, proportionally to their schooling, would double the percentage of returnees (to almost 80% of total migrants). Moreover it would increase the wages of highly educated in Romania and, more importantly, the share of people getting high education. This produces an increase in education and wages by 1-2%. The increase in the schooling-return premium is associated both with a stronger incentive effect that pushes young generations to higher educational achievements, so that more individuals get higher educations, and to a steeper wage profile for highly educated returnees. Certainly, both are a boon for schooling and productivity in the

[9] The simulations reported below are based on our model. The effects on all schooling and wage variables are available upon request.

sending country.

6.4 Discussion

Three interesting implications emerge from our findings with potentially relevant and direct policy implications. The first is that migrants and returnees are strongly affected by wage incentives. Countries with a wage structure strongly increasing in skills attract thus the most highly educated immigrants inducing the greatest benefits to their income and productivity. The second is that migrants acquire valuable skills while aboard and receive significant premia upon return (between 12 and 14% of their wage) relative to non movers. Temporary migration is a reasonable strategy to increase one's lifetime income. Third, for the sending country, under these conditions of selection and return, the average productivity and average human capital, is enhanced when there is a larger share of migrant workers.

One interesting implication of these findings is that in a migration relationship, such as the one described for Romania and the US, it would be beneficial for both countries and for the migrants to allow for temporary working or studying and working visa programs. The strong incentives of the highly educated to migrate and the large premia they would earn upon return make this option very attractive. On the other hand in a migration relationship, such as the one between Romania and Spain, the benefits for both countries may be smaller. Spain receives the less educated workers, still economically useful but less likely to increase productivity and human capital in Spain. Moreover these workers will have less scope for learning and productive increases after return.

A second possibility open to the government of Romania is to provide fiscal incentive, progressive with skills, for people who end their temporary migration period and return[10]. Decreasing the marginal tax proportionally with schooling is like increasing the value of κ, the returns to schooling for people who come back. As shown above, this will increase the incentives of the highly educated to return.

Interesting questions that we are not able to answer with these data is whether the type of selection of returnees will remain the same also for those people who went back to Eastern Europe as a consequence of the great recession (2007-2009) in Western Europe. While the recession in western Europe can be a very important opportunity for Romania and other Eastern European countries to regain some of their expatriates and their enhanced human capital, it is possible that the type of people going back in periods of recession is different from those who returned in periods of normal economic growth. Given the overall positive selection of migrants documented in many studies (e.g. Docquier et al 2010), however, it is likely that returnees, even in recessions, are more educated than non-migrants[11].

Finally the reduction of migration costs (e.g. bureaucratic simplification, skill transferability) can be another lever that European countries may use to increase the beneficial impact of temporary migration. Allowing lower

[10] Such incentive schemes were in fact on the political agenda of the Romanian government but never implemented and abandoned in the context of economic recession.

[11] Moreover, recent evidence suggests that the overall effect of the economic downturn on return migration to Eastern Europe in general and to Romania in particular was rather small.

costs and higher probability of migration and introducing more programs for temporary migration between countries should increase flows. Countries with the right reward structure (compensating highly educated and high quality workers) will attract the best combination of them and the workers themselves will increase their skills and their productivity back home. This is beneficial to the sending countries, to the individuals and to the receiving countries. The losers are countries that are not competitive in attracting the highly educated.

7 Conclusions

In this paper we measure empirically the magnitude and the selection of migrants and returnees in the case of Romania. After 1990 Romania has experienced large emigration flows followed by significant rates of return migration. In terms of countries of destination, magnitude and composition of flows, this pattern of migration is representative for Eastern Europe. Our goal was to characterize the selection of migrants and returnees and to test whether their motivations to migrate and return were consistent with a utility maximizing framework. We used this to assess the effects of migration and return on the level of skills and wages in Romania, allowing also for a education-incentive effect driven by higher expected returns to schooling. Our findings emphasize that return migration is an important phenomenon: about half of the people who migrate do return. Returnees are strongly positively selected, relative to non-migrants, while the selection of migrants depends on the country of destination. Returnees are selected in a similar way as migrants to the countries with highest skill premium (like e.g. the US). Both types of selection (either to migrate or to migrate and return) are consistent with the idea that workers move in accordance to the wage premium they receive. Decisions to return may therefore be part of an optimal strategy to maximize lifetime income. Following the idea that people migrate and return to maximize their utility and that selection at each stage is driven by relative compensation to skills we also perform a simple simulation (based on Peri and Mayr 2009) which suggests that increasing freedom of migration would increase average wage and schooling of the Romanian population through incentives to get educated and the wage-productivity premium to returnees. The overall effects of migration and return on skill and wages of Romanian are positive.

References

Baldwin-Edwards, Martin (2007) Navigating between Scylla and Charybdis: Migration Policies for a Romania within the European Union, *Journal of South-East European and Black Sea Studies*, 7(1), 5 - 35.

Barrett, A and J. Goggin (2010) Returning to the Question of a Wage Premium for Returning migrants, *National Institute Economic Review* 213: R43-R51.

Barro, R. J. and J.W. Lee (2000) International Data on Educational Attainment: Updates and Implications, CID Working Paper No. 42, Cambridge, Ma.

Batista, C., A. Lacuesta and P. Vicente (2010) Testing the 'Brain Gain' Hypothesis: Micro Evidence from Cape Verde, forthcoming in Journal of Development Economics.

Beine, M., F. Docquier and H. Rapoport (2001), Brain Drain and Economic Growth: Theory and Evidence, *Journal of Development Economics* 64 (1), 275-289.

Bhagwati, J. (1976) The Brain-Drain and Taxation: Theory and Empirical Analysis, New York, North Holland.

Bhagwati, J. and C. Rodriguez (1975) Welfare-theoretical Analyses of the Brain-Drain, J*ournal of Development Economics* 2 (3), 195-222.

Bhagwati, J. and K. Hamada (1974) The Brain Drain, International Integration of Markets for Professionals and Unemployment: A Theoretical Analysis, Journal of Development Economics 1 (1), 19-42.

Borjas G. (1987) Self Selection and Earning of Immigrants, *American Economic Review* 77 (4) 531-553.

Borjas, G. and B. Bratsberg (1996) Who Leaves? The Outmigration of the Foreign-Born, *Review of Economics and Statistics*, 78 (1), 165-176.

Budnik, K. B. (2009) Rationality of Post-Accession Migration, *Focus on European Economic Integration* Q1/09: 57-83.

Chiquiar Daniel and Gordon Hanson (2005) International Migration, Self-selection and the distribution of Wages: Evidence from Mexico and The United States, *Journal of Political Economy*, 113(2), 239-281.

Clemens M., C. Montenegro and L. Pritchett (2008) The Place Premium: Wage Differences for Identical Workers across the US border, Center for Global Development Working Paper # 148, July 2008.

Co, C.Y., I. Gang and M. Yun (2000) Returns to Returning, *Journal of Population Economics* 13: 57-79.

Commander, S., R. Chanda, M. Kangasniemi and A. Winters (2008) The Consequences of Globalisation: India's Software Industry and Cross-Border Labour Mobility, World Economy, 31, 2: 187-211.

De Coulon, A. and M. Piracha (2005) Self-Selection and the Performance of Return Migrants: the Source Country Perspective, *Journal of Population Economics*, 18, 779-807.

De Coulon, A., and J. Wadsworth (2010): On the relative rewards to immigration: a comparison of the relative labour market position of Indians in the USA, the UK and India, *Rev Econ Household* 8:147–169.

Diminescu, D. (ed.) (2003) Visibles mais peu nombreux. Les circulations migratoires roumaines, Editions de la Maison des Sciences de l'Homme, Paris.

Diminescu, D.and S. Lazaroiu (2002) Circular migration of Romanians after 1989: migrants behavior, institutions and policies, IOM-Study, Bucharest.

Docquier, F. and A. Marfouk (2006) Measuring International Migration by Educational Attainment, 1990-2000, in: C. Ozden and M. Schiff (eds.), International Migration, Remittances and the Brain Drain, Chapter 5, 151-199, Palgrave McMillan, New York.

Docquier, F. and H. Rapoport (2008) Brain Drain and Human Capital Formation in Developing countries: Winners and losers, The Economic Journal 118 (528), 631-652.

Dustmann, C. (1995) Savings Behavior of Migrant Workers - A Life Cycle Analysis, *Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 4: 511-533.

Dustmann, C. and A. Glitz (2011) Migration and Education, *Handbook of the Economics of Education*, Vol. 4 (forthcoming).

Dustmann, C. and Y. Weiss (2007) Return Migration: Theory and Empirical Evidence from the U.K., *British Journal of Industrial Relations*, 45 (2), 236-256.

Earle J.S. and C. Pauna (1996) Incidence and duration of unemployment in Romania, *European Economic Review* 40: 829-837.

Earle J.S. and C. Pauna (1998) Long-term unemployment, social assistance and labor market policies in Romania, *Empirical Economics* 23: 203-235.

EIR (2003) Pre-Accession Impact Study on Migration and Free Movement of Persons, European Institute of Romania, Bucharest.

Epstein, G. and D. Radu (2007) Returns to Return Migration and determinants of Subsequent Moves, Manuscript, University of Edinburgh, February 2007.

Erlick, T. and O. Ciobanu (2009) Migration networks and policy impacts: insights from Romanian–Spanish migrations, *Global Networks* 9 (1): 100-116.

EU-SILC (2004) European Union Statistics on Income and Living Conditions" Eurostat, 2009 available at http://epp.eurostat.ec.europa.eu/portal/page/portal/microdata/eu_silc

Fernandez Huertas-Moraga, Jesus (2008) New Evidence on Immigrant Selection, Manuscript Universitat Autonoma, April 2008.

Grogger J. and G. Hanson (2008) Income Maximization and the Strength and Sorting of International Migration, NBER Working Paper No. 13821, Cambridge, Ma.

Grubel H. and A. Scott (1966), The International Flow of Human Capital, *American Economic Review* 56 (1/2), 268-274.

Hartog, J. and R. Winkelmann (2003): Comparing migrants to non-migrants: The case of Dutch migration to New Zealand, *Journal of Population Economics* 16:683–705.

Hazans, M. (2008) Post-enlargement return migrants' earnings premium: Evidence from Latvia, EALE 2008 paper No. 541.

Iara, A. (2008): Skill Diffusion by Temporary Migration? Returns to Western European Working Experience in the EU Accession Countries, WIIW Working Paper 46, July 2008, Vienna.

Kaestner R. and O. Malamud (2010) Self Selection and International Migration: New Evidence from Mexico, NBER Working Paper # 15765, Cambridge, Ma.

Kahanec, M. and K. F. Zimmerman, eds. (2010) EU Labor Markets after Post-Enlargement Migration, Berlin: Springer.

Kilic, T, C. Carletto, B. Davis and A. Zezza (2009): Investing back home: Return migration and business ownership in Albania, *Economics of Transition* 17(3): 587–623.

Luo Y.L. and W.J. Wang (2002) High-skilled Migration and Chinese Taipei's Industrial Development in: International Mobility and the Highly Skilled, OECD editions, Paris.

Martin, R. And D. Radu (2011) Return Migration: The Experience of Eastern Europe, *International Migration* (forthcoming).

Mayr K. and G. Peri (2009) Brain Drain and Brain Return: Theory and Application to Eastern-Western Europe" in *Berkeley Electronic Journal of Economic Analysis & Policy*, Contributions. Vol. 9: Issue 1, Article 49.

McCormick B. and J. Wahba (2001), Overseas Work Experience, Savings and Entrepreneurship Amongst Return Migrants to LDC's, *Scottish Journal of Political Economy* 48 (2), 164-178.

OECD (2008): International Migration Outlook: SOPEMI, Paris.

Ortega F. and G. Peri (2009) The Causes and Effects of International Migrations: Evidence from OECD Countries 1980-2005 NBER Working Paper #14833, April 2009.

Piracha M. and F.Vadean (2010) Return Migration and Occupational Choice: Evidence from Albania, *World Development* 38 (8): 1141-1155.

Radu, D. (2003) Human Capital Content and Selectivity of Romanian Emigration, Discussion Paper 365, Hamburg Institute of International Economics.

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

Sandu D. et al. (2006) Living abroad on a Temporary Basis. The Economic Migration of Romanians: 1990-2006, Open Society Foundation, Bucharest.

Santos, M. and F. Postel-Vinay (2003) Migration as a source of Growth: The Perspective of a Developing Country, *Journal of Population Economics* 16 (1): 161-175.

Şerban, M. and B. Voicu (2010): Romanian migrants to Spain : In- or outside the migration networks – A matter of time?, *Revue d'études comparatives Est-Ouest*, 41 (4): 97-124.

United Nations (2009) Trends in International Migrants Stocks: The 2008 Revision, Department of Economic and Social Affairs, The United Nations Population Division. May 2009.

Voicu, A. (2005) Employment Dynamics in the Romanian Labor Market: A Markov Chain Monte Carlo Approach, *Journal of Comparative Economics* 33(3): 604-639.

Zucker L. and M. Darby (2007) Star Scientists, Innovation and Regional and National Migration, NBER Working Paper No. 13547, Cambridge, Ma.

Tables

Table 1
Stock of Emigrants to OECD countries as ratio of the population in the home country

Country of origin	1990	2000
Albania	0.162	0.190
Bulgaria	0.060	0.080
Croatia	0.123	0.140
Czech Republic	0.021	0.027
Estonia	0.028	0.054
Hungary	0.042	0.041
Latvia	0.020	0.033
Lithuania	0.057	0.054
Macedonia	0.138	0.169
Poland	0.041	0.044
Romania	0.020	0.031
Russia	0.003	0.006
Serbia and Montenegro	0.069	0.091
Slovenia	0.044	0.072

Source: Docquier and Marfouk (2006). The data are obtained by national censuses of the receiving country and include all OECD countries as receiving. The data are relative to the census year of the receiving countries and those are clustered around 1990-1991 and 2000-2001.

Table 2

Imputed return relative to gross migration flows (any OECD destination), 1990-2000;

Selected Eastern European source countries

Source	Return Flows (imputed)	Gross flows	Return/Gross
Albania	20476	34207	0.60
Bulgaria	24353	42109	0.58
Czechoslovakia	24230	18697.5	1.30
Estonia	5859	12099	0.48
Hungary	54450	40535	1.34
Lithuania	2824	12010	0.24
Latvia	3053	9713.5	0.31
Poland	282984	306841.5	0.92
Romania	54197	132311.5	0.41

Note: The data are obtained by authors' calculations (as described in the text) using Docquier and Marfouk (2006) and the UN (2009) datasets. The return flows are imputed assuming that all re-migration is return migration.

Table 3

Imputed return relative to gross emigration flows from Eastern Europe, 1990-2000;

Selected OECD destination countries

Destination	Return Flows (imputed)	Gross flows	Return/Gross
Australia	28933	15012	1.93
Austria	26385	110096	0.24
Belgium	11219	13151	0.85
Canada	101096	108537	0.93
Finland	1007	9265	0.11
France	49413	19982	2.47
Norway	4247	6649	0.64
Sweden	19483	22684	0.86
USA	230643	303148	0.76

Note: The data are obtained by authors calculations (as described in the text) using Docquier and Marfouk (2006) and the UN (2009) datasets. The return flows are imputed assuming that all re-migration is return migration.

Table 4
Romania: Migrants and returnees, by education (from aggregate and NDS data)

Sample:	Romania NDS, 2003	OECD Country Census2001
	returnee as ratio of population	living abroad (OECD) as ratio of population
all	0.049	0.032
education groups		
tertiary	0.058	0.126
secondary	0.056	0.126
primary completed	0.034	0.016
No degree completed	0.015	0.039

Note: Authors' calculations on NDS 2003 data and Docquier and Marfouk (2006) data.

Table 5
Main Countries of Destinations for Romanian Migrants

Country of Destination	Migrants as share of Romanian migrated to OECD countries
USA	0.230
Germany	0.164
Italy	0.113
Canada	0.102
Spain	0.083
Austria	0.068

Note: Authors' calculations on Docquier and Marfouk (2006) data. The sum of Romanian migrants to OECD countries is standardized to 1. The table reports the fraction of such sum in each of the top 6 countries of destination.

Table 6

Average Selection on Observable Skills, relative to Non Migrants

Romania, 2003

	(1)	(2)	(3)	(4)	(5)
	Average Selection on Observable Skills Census 2002	**Average Selection on Skills NDS 2003**	**Average selection Correcting for Participation Census 2002**	**Average selection of Men Census 2002**	**Average selection on Observables Wage-earning abilities (Excluding Ethnic minorities)**
Returnees	+0.14	+0.12	+0.13	+0.13	+0.14
Migrants to US	+0.16	+0.13	+0.20	+0.14	+0.15
Migrants to Austria	0.03	0.01	+0.04	+0.03	0.02
Migrants to Spain	-0.11	-0.13	-0.07	-0.13	-0.10

Note: The calculation of Average selection on Observable Skills follows the formulas in section 3.2 of the text. Column (1) uses the employment data by skill cell from the Romanian Census 2002, Column (2) uses the employment data from the National Demographic survey 2003, Column (3) corrects for participation in Romania; Column (4) includes only male individuals. Specification (5) excludes the Roma ethnic group among non migrants.

Table 7:

Migration- and Return- Premium and their correlation with wage-earning skills

Linear monthly Wage premium

		(1)	(2)	(3)	(4)
Dependent Variable		**Return Premium**	**Migration to US Premium**	**Migration to Austria Premium**	**Migration to Spain Premium**
Explanatory variables	**Specification:**				
Ln(Wage Non Movers)	Basic	0.36**	1.14**	-0.04	-0.75**
		(0.01)	(0.04)	(0.03)	(0.04)
	With Control for age effects	0.49**	1.28**	0.14**	-0.44**
		(0.02)	(0.04)	(0.04)	(0.03)
	With Control for Family structure	0.40**	1.16**	-0.01	-0.65**
		(0.03)	(0.04)	(0.02)	(0.03)
Education Dummies (Reference Group is no Education)	Primary Completed	0.10**	0.92**	-0.04	0.35*
		(0.01)	(0.02)	(0.04)	(0.03)
	Secondary Completed	0.02	0.57**	0.12**	0.13**
		(0.01)	(0.03)	(0.02)	(0.02)
	Tertiary Completed	0.40**	0.93**	0.99**	0.18**
		(0.02)	(0.04)	(0.03)	(0.02)
Type of Skill Premium		Positive Skill-Premium	Positive Skill-Premium	Neutral Skill-Premium	Negative Skill-Premium

Note: The unit of observation is an education-age-gender-family status cell. There are 320 cells. The dependent variable is the difference between the wage of a returnee (migrant) and that of a non-mover in the same skill cell expressed in thousands of 2003 $. Method of estimates is weighted LS, with weights equal to the non-migrant population in the cell.

Table 8

Correlation between Migration- and Return- frequencies and migration and return premium

Linear monthly Wage premium

		(1)	(2)	(3)	(4)
Dependent Variable		Relative frequency of return	Relative frequency of migration to US	Relative frequency of migration to Austria	Relative frequency of migration to Spain
Explanatory					
Premium	In Population cells	0.38** (0.05)	0.24** (0.01)	0.30** (0.02)	0.63** (0.04)
	In Employment Cells	0.21** (0.02)	0.27** (0.03)	0.18** (0.03)	0.27** (0.05)
Premium	Controlling for age and family effects	0.11** (0.02)	0.33** (0.02)	0.15** (0.03)	0.01 (0.02)
Education Dummies (Reference Group is no Education)	Primary Completed	0.38** (0.03)	0.28** (0.03)	1.01** (0.03)	1.02* (0.03)
	Secondary Completed	0.82** (0.02)	0.93** (0.02)	1.19** (0.03)	-0.02 (0.02)
	Tertiary Completed	0.80** (0.02)	4.42** (0.02)	1.20** (0.03)	-0.13** (0.03)
Does it support the selection by premium theory?		Yes	Yes	Yes, in part	Yes in part

Note: *The unit of observation is an education-age-gender-family status cell. There are 320 cells. The explanatory variable is the difference between the wage of a returnee (migrant) and that of a non-mover in the same skill cell expressed in thousands of 2003 $. Method of estimates is weighted LS, with weights equal to the non-migrant population in the cell unless otherwise noted.*

Table 9

Simulated effects of increasing freedom of migration, p, in Romania on schooling and wages

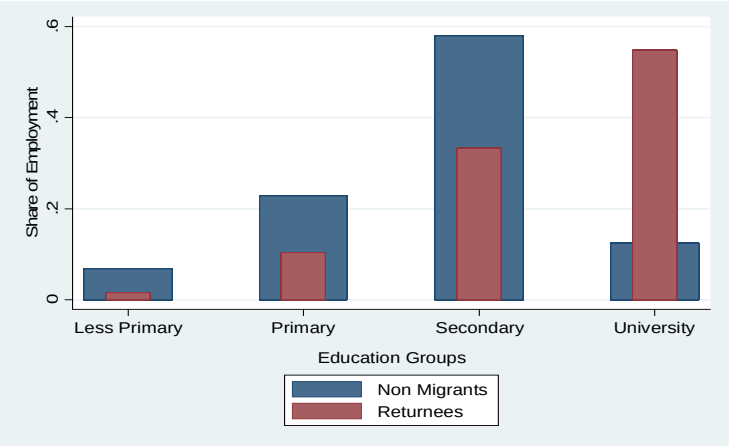
Parameter as in Romania 2003; Migration and return

p	**0**	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**
Years of Schooling							
Average Schooling of young	12	12.16	12.33	12.49	12.64	12.78	12.91
Average schooling of old	12	12.32	12.66	13.01	13.37	13.75	14.15
Average schooling, overall	12	12.24	12.50	12.76	13.03	13.30	13.59
Wages (standardized to 1 with no migration)							
Average wages, young	1	1.01	1.02	1.04	1.05	1.07	1.08
Average wages, old	1	1.04	1.09	1.15	1.22	1.29	1.36
Average wages, overall	1	1.03	1.06	1.10	1.14	1.18	1.23
Average wage No primary	0.47	0.47	0.47	0.47	0.47	0.47	0.47
Average wage Primary-Secondary	0.73	0.73	0.73	0.73	0.73	0.73	0.73
Average wage tertiary-young	1.41	1.43	1.45	1.48	1.50	1.52	1.55
Average wage tertiary-old	1.41	1.48	1.56	1.64	1.73	1.82	1.92
Migration Rates							
Share of emigrants	0	0.045	0.091	0.137	0.183	0.230	0.276
Share of returnees among emigrants		0.471	0.482	0.492	0.502	0.512	0.521

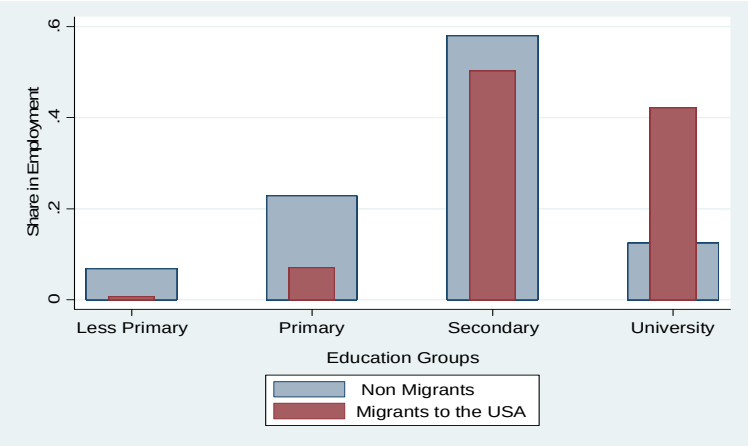
Note: We standardized all the wages to be relative to the average wage in the case of no emigration. The simulations follow Mayr and Peri (2009).

Figure 1: Selection over Education

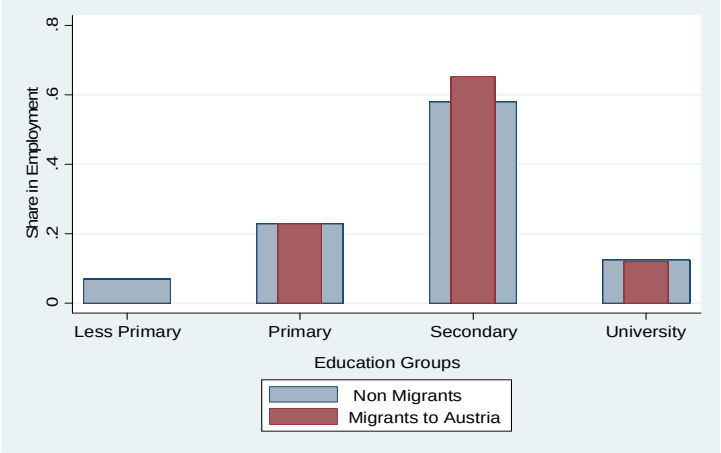
Panel 1: Non Migrants and Returnees (Romania Census)



Panel 2: Non Migrants and Migrants to USA (US Census)



Panel 3: Non Migrants and Migrants to Austria (Austria census)



Panel 4: Non Migrants and Migrants to Spain (Spanish Census)

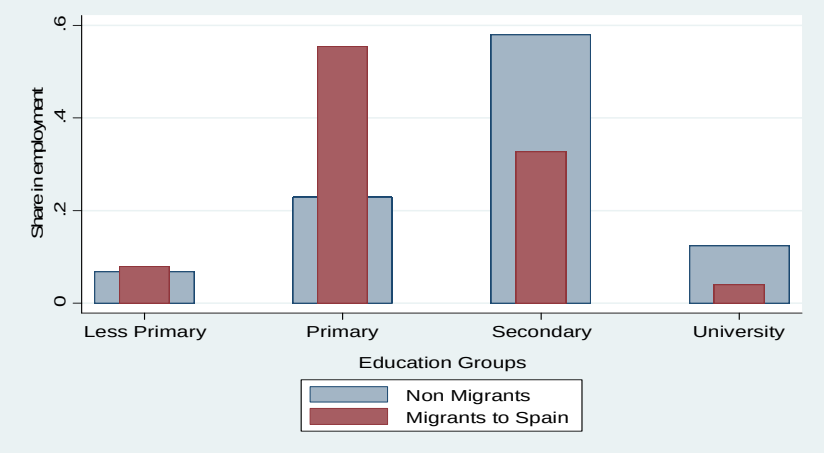
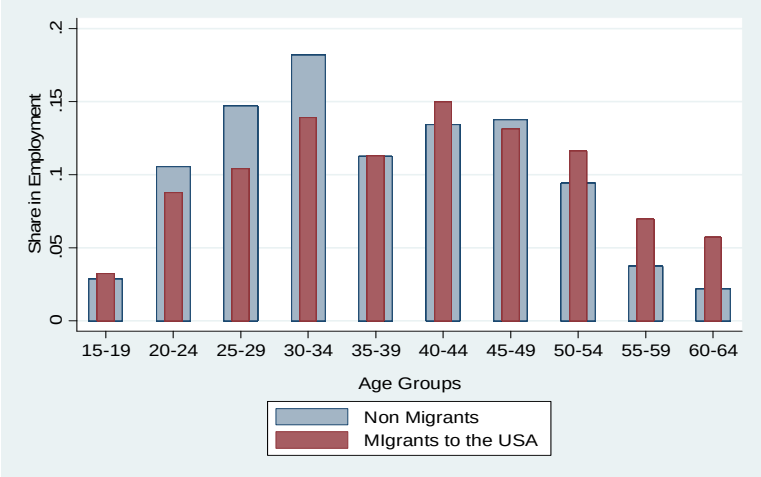


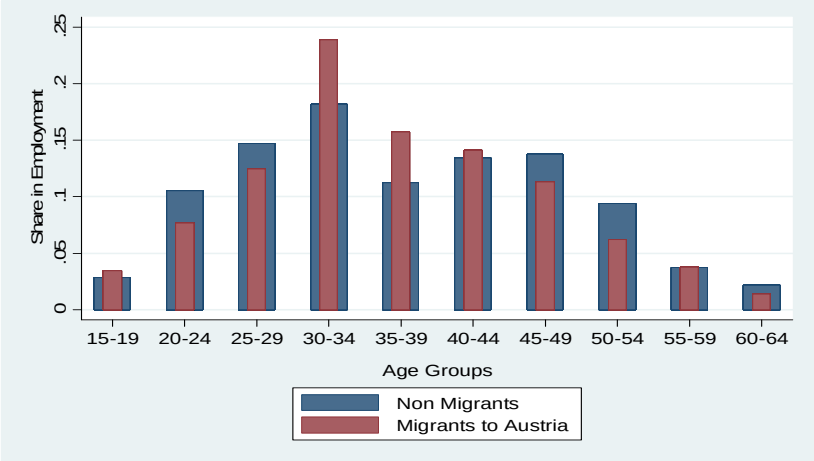
Figure 2: Selection over Age

Panel 1: Non Migrants and Returnees(Romania Census)

Panel 2: Non Migrants and Migrants to USA (US census)



Panel 3: Non Migrants and Migrants to Austria (Austria census)



Panel 4: Non Migrants and Migrants to Spain (Spanish Census)

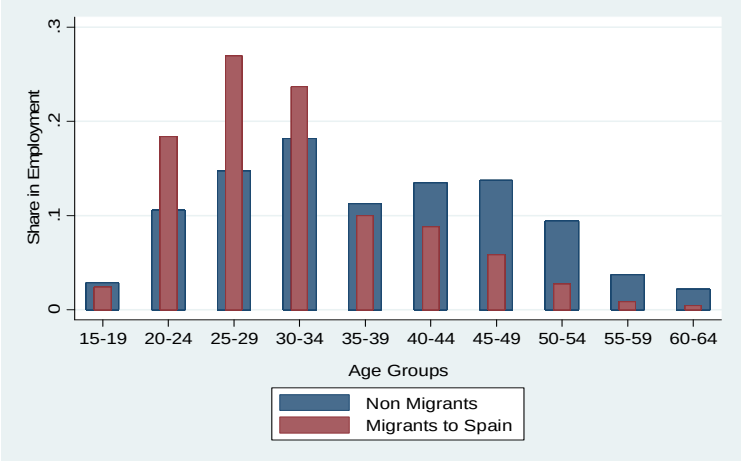
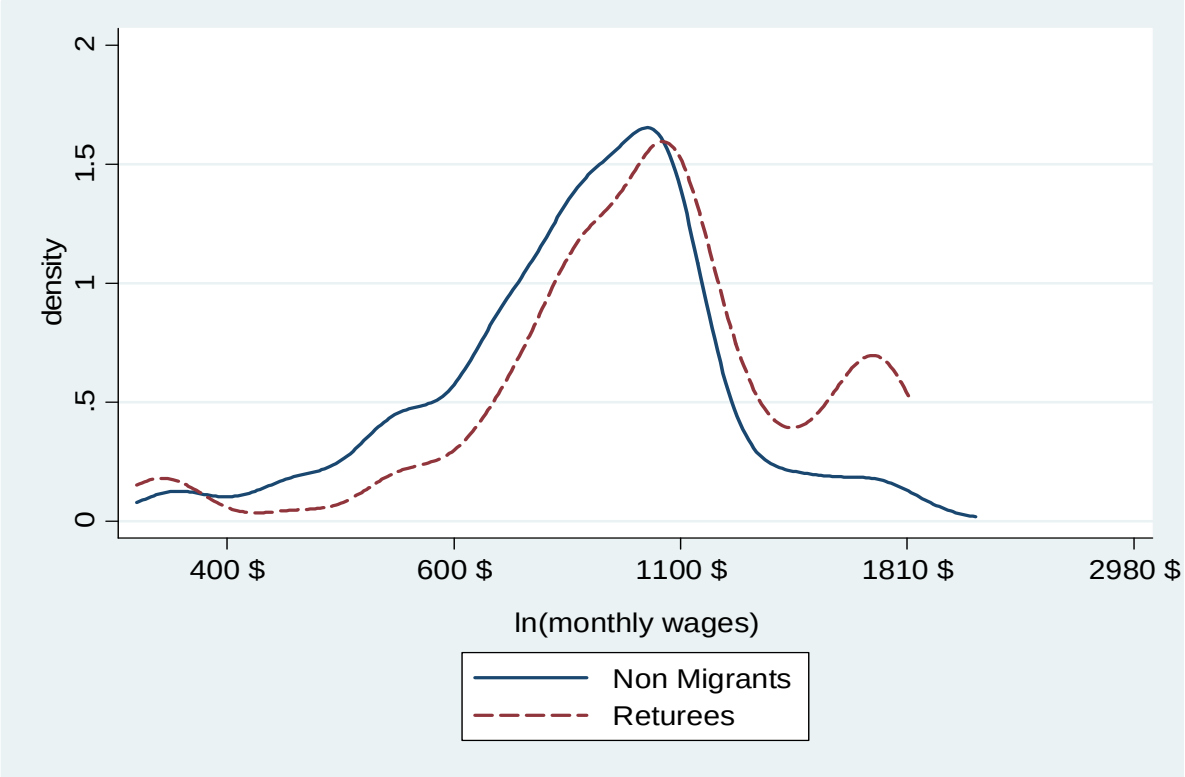


Figure 3

Kernel density of non migrants and returnees over skill, Census 2002

Monthly wages, in 2003 US $



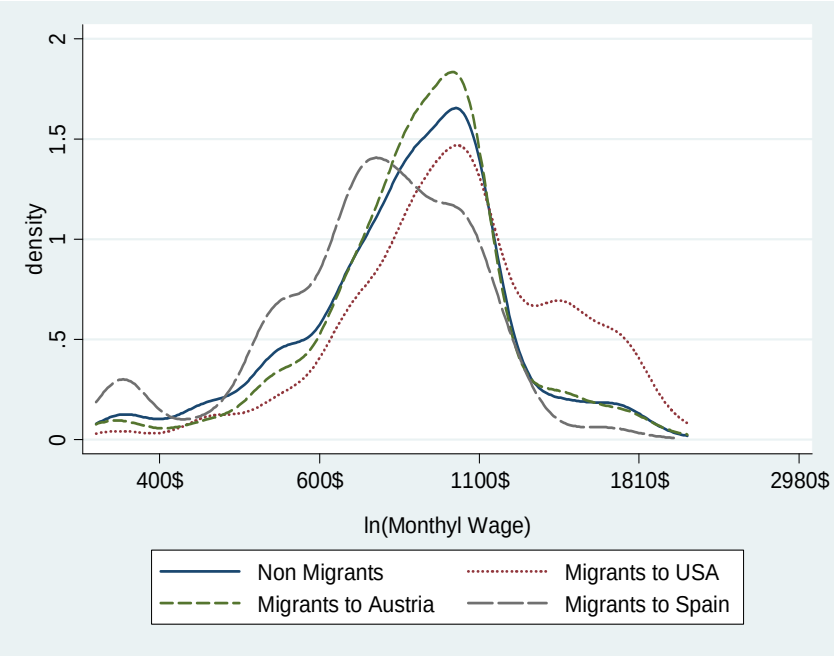
Note: The function represents the density of each population over the wage-earning skill, ln(wage), distribution estimated using a Gaussian kernel. Bandwidth is chosen optimally, for each distribution, following Fernandez-Huertas (2008).

Figure 4

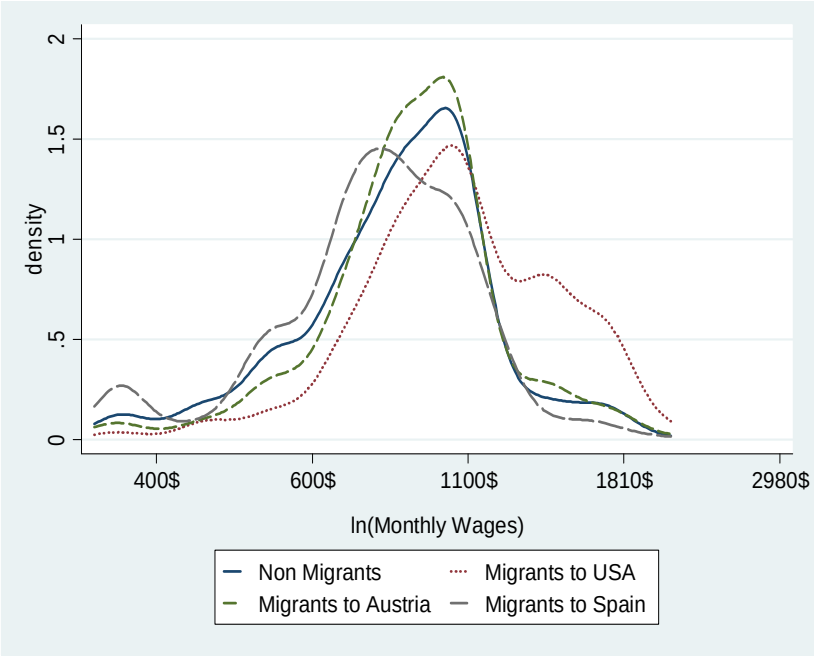
Kernel density of migrants and returnees over skill, Census 2002

Monthly wages, in 2003 US $

Panel 1: Based on employment



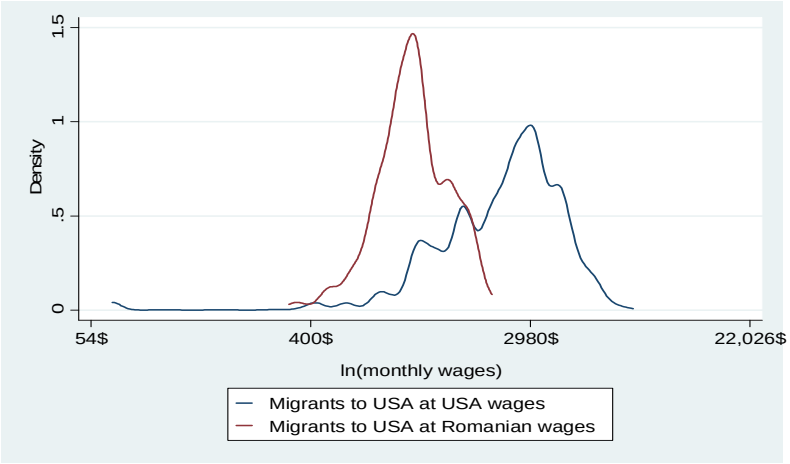
Panel 2: Based on Population



Note: The function represents the density of each population over the wage-earning skill, ln(wage), distribution estimated using a Gaussian kernel. Bandwidth is chosen optimally for each distribution, following Fernandez-Huertas (2008).

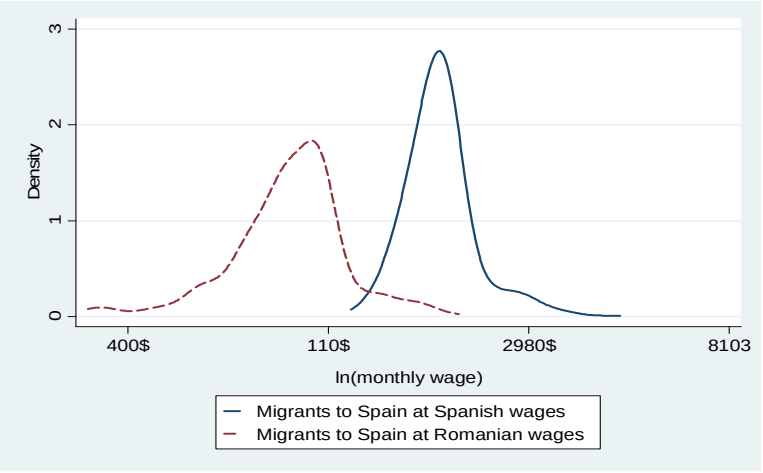
Figure 5: Migration premium: Wages in the Destination country and in Romania

Migrants to US



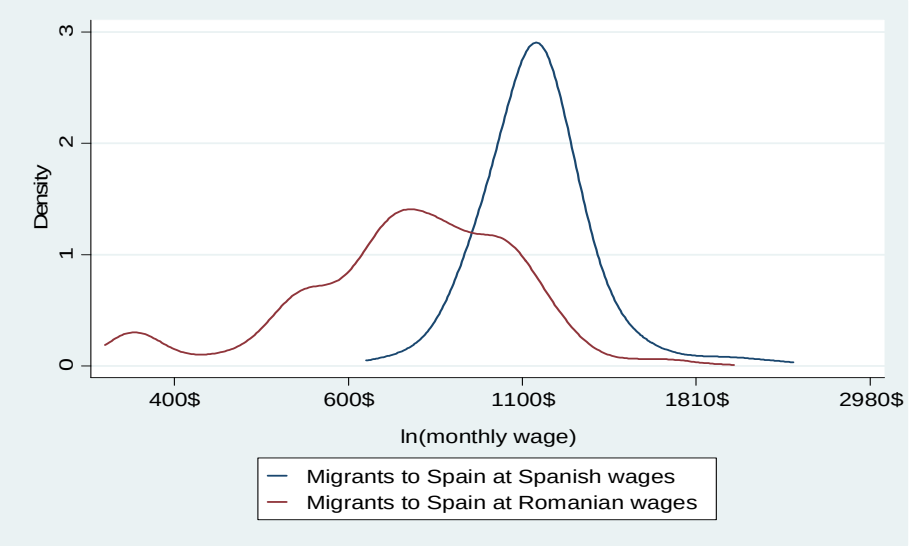
Average premium = 112% of Romanian wage =990 $ per month

Migrants to Austria



Average premium=100% of Romanian wage =882 $ per month

Migrants to Spain



Average premium=34% of Romanian wage =300 $ per month