

Self-Selection and the Performance of Return Migrants: A Source Country Perspective*

Augustin de Coulon[†] and Matloob Piracha[‡]

April 22, 2004

Abstract

In this paper, using the framework of a Roy theoretical model, we examine the performance of return migrants in Albania. We ask two main questions: (i) Had they chosen not to migrate, what would be the performance of return migrants compared to the non-migrants? and (ii) What would be the performance of non-migrants had they decided to migrate and return? Both the selection estimates and the semi-parametric approach allow us to conclude that the flows of return migrants are negatively selected. We find that, had they decided to migrate and come back, the non-migrants would have earned more than twice the wages of return migrants.

JEL CLASSIFICATION: F22, J61

KEYWORDS: self-selection, return migrants, Albania

*We would like to thank Alan Carruth, Arnaud Chevalier, Ira Gang, Alan Manning, Roger Vickerman, two anonymous referees and participants at the Applied Econometrics Association conference, Bruxelles, Centre for Economic Performance Labour Market Workshops, London, and the 18th Annual Congress of European Economic Association, Stockholm, for helpful comments on earlier versions of the paper. This work is part of the project "Borders, Migration and Labour Market Dynamics in a Changing Europe" financed by the UK Economic and Social Research Council, Grant No L213251042. The usual disclaimer applies.

[†]Queen Mary, University of London and Centre for Economic Performance, London School of Economics, Houghton St, WC2A 2AE London, UK. (Fax: ++44-207-9557595; e-mail: a.de-coulon@lse.ac.uk)

[‡]University of Kent, Keynes College, Canterbury, Kent CT2 7NP, UK. (Fax: +44- 1227-827850; e-mail: M.E.Piracha@kent.ac.uk)

1 Introduction

International migration is a selective process. Human capital models of migration claim that those who choose to leave a country might be more able and/or more motivated than those who choose to stay in their home country (see Chiswick, 1999). If this is the case then immigrants are said to be positively selected compared to the home population. Borjas(1987, 1991), however, showed that immigrants coming from a country with more unequal wage distribution than that prevailing in the host country may be negatively selected. In an extension of this work, Borjas and Bratsberg (1996) investigated the return migration of foreign-born individuals in the United States and showed how this may influence the type of self-selection characterising the migration flows. Dustmann (2003) studies the optimal length of stay abroad and return behaviour of temporary migrants in the framework of life-cycle analysis while Dustmann and Kirchkamp (2002) look at the activity choice of return-migrants. Bauer et al. (2002), studying Portuguese immigrants in Germany, conclude that the German guest worker system succeeded in attracting positively self-selected immigrants in terms of unobservable characteristics and compared to the native German workers. Chiquiar and Hanson (2002) study the performance of Mexican immigrants in the United States and compare them to the non-migrant Mexicans. Using the semi-parametric approach of DiNardo et al.(1996), they reject previous results found in a more descriptive literature that Mexican immigrants in the US tend to be negatively selected in terms of observable skills compared to the stayers.

Most of the literature on return migration deals with the issue of self-selection within the context of the host country. One exception, however, is Co et al. (2000), who study the potential economic benefits generated by the returning Hungarian migrants after spending time abroad. They address two potential selection biases: one due to the decision to migrate vs. to stay home, and the second due to the decision to work vs. not to work. Using the maximum likelihood framework they find negative selection for women but not for men.

The focus of this paper is similar to that of Co et al., however, unlike them, we study the wage effects of return migration in Albania, comparing the performance of returnees to those who stayed in the home country. More specifically, we address the question of the self-selection process of out- and then re-migration of the individuals who left the source country and then returned home using the stayers (non-migrants) as the counter-factual. We address the following questions: (i) Had they chosen not to migrate, what would be the performance of return-migrants compared to those who stayed? and (ii) What would be the performance of non-migrants had they decided to migrate and return?¹ To answer these questions, we use a sample

¹Note that our focus (treated) group is those who migrated and then returned to Albania and therefore we do not correct

of 694 Albanian individuals and use two alternative methodologies, a selection model along the lines of Heckman (1979) and Lee (1982) and a semi-parametric approach proposed by DiNardo et al. (1996). The first approach allows us to directly address the questions but offers only mean conditional earnings, while using the second approach we can study the effect of migration on the entire wage distribution.

Evidence suggests that a large number of migrants from Central and Eastern European Countries fall into the category of temporary (or guest) workers. For example in Greece, amongst the Albanians who received a temporary white card in the regularisation programme in 1998, only 54% proceeded to the second phase of application one year later to obtain a permanent green card. In a survey realised in Albania by the International Organisation for Migration (IOM) in 1992, 79% of respondents said they were ‘likely’ or ‘very likely’ to migrate for a few months and only 24% wanted to settle permanently in another country (IOM, 1995). Other evidence based on Eurobarometer shows that 50% of Albanians planned to emigrate for a short period only (see Papapanagos and Sanfey, 1998).

Our paper adds to the limited literature on the analysis of return-migrants in their home labour market in the context of a self-selection model. This is the first study of such an issue for Albania, a transition country most affected by migration. Furthermore, this is the first paper to use a semi-parametric kernel density approach to study the impact of return migration.

We find support for the negative self-selection of return-migrants compared to the native non-migrant population (stayers). Our empirical results show that stayers would have performed much better than return-migrants had they chosen to migrate. We argue that for stayers the decision not to migrate comes from the non-transferability of current skills due to language barrier, and also by the low added return to human capital in the host country. Interpreting those results in the framework of our model, we find support to a story of negative selection of the wave of return-migrants compared to non-migrants. These results have potential implications for migration policies of the host and the source countries.²

The rest of the paper is organised as follows. Brief background on Albanian migration is presented in Section 2, while the theoretical model is discussed in Section 3. Section 4 describes the data set and selection of the variables. In Section 5, empirical methodology used to examine the issues raised in the theoretical model is presented while the empirical results are given and discussed in Section 6. Concluding remarks and

for people who went abroad and didn’t return. We are therefore relying on the assumption that return-migrants are randomly selected from the pool of immigrants. This appears to be the case for at least one major EU host country, Germany (see Constant and Massey, 2003).

²We would like to emphasise that "return-migrants" in our paper are only those who participated in the labour market upon their return to the source country after spending some time abroad. Therefore, all policy implications of our results should be considered with that in mind.

potential policy implications appear in the last section.

2 Albanian migration: A brief background

Perhaps because of its central location in Europe and its relative poverty, Albania has long been a country of emigration. However, between 1945 and 1990 the state pursued a policy of social and economic isolation, totally restricting any movement of its citizens out of its borders. During the transition period, a large number of people, uncertain about the economic prospects of Albania, left the country. This was taking place against the backdrop of rapid and radical political change that had already begun elsewhere in Central and East European countries (CEEC) at the end of the 1980s. These events provided a further catalyst for change in Albania and helped to put in motion the organisational skills and energy of those who had been waiting for the right time to leave.

Precise figures on Albanian immigrants are difficult to gather due to the potentially high number of non-declared (illegal) individuals either settled or working short time periods in the host countries. For example, officially 4300 Albanians were issued a residence permit in 1997 in Greece. But when the country adopted a regularisation programme (between November 1997 and May 1998) for undocumented immigrants, 239,000 Albanian immigrants applied (see SOPEMI, 2000). Hence, behind the official figures, there are a rather large number of undocumented migrants not only in Greece but elsewhere in Europe also, particularly in Italy. The Albanian Center for Economic Research (2002) estimates that at least 15% of the Albanian population is living abroad, which is by far the highest proportion amongst the Central and East European economies.

A gradual improvement of the economic situation of Albania took place until mid-1996, owing mainly to remittances and macroeconomic policies³. These factors lessened, to a certain extent, the major economic and social problems, which emerged as a result of high unemployment rates and big disparities in wealth. However, these “positive factors” proved temporary as the domestically financed deficit increased to almost 11 percent of GDP, and inflation tripled to more than 17 percent by the end of 1996. This was exacerbated by the collapse of the pyramid schemes in early 1997, causing an estimated loss of savings of about \$1 billion⁴.

The worsening economic situation led to a second large outflow of individuals as employment prospects in Albania dwindled for many. Emigration has an important impact in the reduction of unemployment in

³Remittances have played a key role in the development process of not only Albania but other CEECs also. See Leon-Ledesma and Piracha (2001) for an analysis of the role of remittances in selected CEECs.

⁴Pyramid schemes were companies that, by claiming to be engaged in profitable investments, attracted large and increasing volumes of funds from private depositors with promises of dramatically high returns. In reality, however, depositors' funds were largely not used for "growth generating" investments, but served either to pay interest on existing deposits or were transferred by the schemes' owners to bank accounts abroad. For a detailed analysis of the pyramid scheme crisis see Jarvis (2000).

the country. According to official data, during 1998 unemployment in the country reached 17.7%, with a figure of 19.1% in the north-eastern areas where the level of emigration is lower and 13.4% in the south where mass emigration exists. Given that Albanian emigration is often driven by seasonal and temporary employment, this has had an impact on the Albanian labour market. It is estimated that half the overall number of emigrants are seasonally employed in the host countries.

According to data from the Albanian Ministry of Labour and Social Affairs, during the last ten years, Albanians have emigrated to about twenty European countries. However by far the largest number goes to Greece followed by Italy. It may be the result of easier access to information about job availability and level of wages in Greece, and also relatively lower transportation costs. The migration flow is amplified by the need for a flexible non-unionised workforce for the informal economy in Greece. However, as mentioned before, most of the migration appears to be temporary and for a specific purpose: to raise funds to setup enterprise in Albania and/or to acquire skills by working in a relatively richer and established market economy.

3 Theoretical framework

In earlier literature, migration has been modeled as a one shot move, where individuals take their decision following an income maximising strategy to either migrate or stay in the home country (Harris and Todaro, 1970). More recently, migration has been considered as a dynamic process within the lifetime expectations of workers (Djajic, 1988; Dustmann, 1997). In this context, there is evidence that migration is self-selective, i.e., those who migrated would have done better regardless of whether or not they had gone abroad. Immigrants are often found to be “more able and more highly motivated” than those who stay at home. In this study we question this assertion. We analyse the performance of return-migrants in the source country, i.e., those who migrated but then decided to return to participate in the labour market of the source country⁵.

Using Albanian data, we want to know if migrants who returned home to Albania were selected from the upper or lower part of the ability distribution in the source country. To conduct such an analysis we investigate their performance once they return to Albania. The problem can be modelled by assuming income maximising individuals who make a migration decision based on their expected income in the source and the host countries net of any migration (and remigration) costs. More formally, we use a version of the Roy (1951) theoretical model modified by Borjas (1987, 1999) and Borjas and Bratsberg (1996) to analyse this problem. But in contrast with those papers, we analyse the impact of self-selection on the home country

⁵We ignore the individuals who return to spend their retired life in the source country.

rather than the host country.

Let the log earnings distribution in the source country be,

$$w^s = \mu^s + \eta\nu \tag{1}$$

where μ^s is the mean of log income in the source country, η is interpreted as the rate of return to skills in the source country relative to that in the host country and is assumed to be known to the migrant and ν is the random variable that measures deviations from the mean and is independently and normally distributed with mean zero and variance σ_ν^2 . Now let the log earnings facing the population of the source country when they decide to migrate to the host country be,

$$w^h = \mu^h + \nu + \epsilon \tag{2}$$

where μ^h is the mean income that migrants receive in the host country and ϵ is the random variable that measures deviations from the mean income in the host country and is not known to the migrant, i.e., it captures the luck and/or misinformation about the prospects in the host country, it is assumed to be independently and normally distributed with mean zero and variance σ_ϵ^2 .

One of the main reasons for migration from Albania to EU countries is the significant wage gap between the two countries. A temporary migration to Western Europe (primarily to Greece and Italy) offers higher paid employment and the potential to acquire skills, and, moreover, helps overcome any capital constraints that an individual may face in the source country to start an enterprise.⁶ Therefore, migrants will only incur migration costs if they expect that after spending a fraction δ of their working life in the host country they can increase their earnings by some percent, κ , when they return to their home country. We assume that the parameters δ and κ are constant.

Workers in Albania, therefore, have the following option: residing in an EU country for a fraction of the working life, followed by a permanent return to the source country. Ignoring discounting and using a first-order approximation, the log earnings associated with this choice (w^r) are given by:

$$w^r = \delta w^h + (1 - \delta)(w^s + \kappa) \tag{3}$$

⁶Mesnard (2000) analyses the choice of activity of return migrants taking into account credit constraints in the home country.

where δ and κ are parameters as defined above.

Workers maximise their lifetime earnings net of all migration costs. For the migration motive to be relevant, a person will only migrate if the expected earnings (due to skill acquisition abroad) in the source country, after returning, are greater than earnings in the source country if the individual did not migrate, net of both migration and remigration costs. Formally, we can write this as:

$$Ew^r > w^s + C^m + C^r \quad (4)$$

where C^m and C^r are the migration and remigration costs respectively.⁷

Substituting eqs (1),(2) and (3) in (4), we get the condition under which a person will migrate (with the intention of returning to the source country).

$$(1 - \eta)\nu > (\mu^s - \mu^h + \kappa) + \frac{C^m + C^r - \kappa}{\delta} \quad (5)$$

Note that so far we have been assuming that a migrant must return to Albania as he is either required to or has already decided at the time of migration to return home. However, to complete the picture, it could be the case that the migrant could stay, either permanently or for a relatively longer period of time, in the host country.⁸ In this circumstance, we need to set out the conditions under which (i) a person will migrate regardless of future intentions and (ii) once migrated, the person will return to the source country after spending a fraction of time in the host country, i.e., has no incentive to stay in the host country permanently. The two conditions are respectively given as,

$$Ew^h > w^s + C^m \quad (6)$$

and

$$Ew^r > w^h + C^r \quad (7)$$

Eq (6) states that if the expected wage net of migration costs is greater than the wage in the source

⁷This include both pecuniary and non-pecuniary costs of migration.

⁸As was discussed in Section 2, there are some who successfully applied for a permanent stay in Greece.

country then it is better for a person to migrate. However, once abroad, a migrant will return to the source country only if the expected wage upon return, net of remigration costs, is greater than the wage in the host country. Substituting for the wages from the above equations, we get the following conditions under which a person will migrate regardless of future intentions

$$(1 - \eta) \nu > \mu^s - \mu^h - C^m \quad (8)$$

and will migrate and then return home after spending a fraction of time in the host country,

$$(1 - \eta) \nu < (\mu^s - \mu^h + \kappa) - \frac{C^r + C^m - \epsilon}{1 - \delta} \quad (9)$$

It is easier to explain the intuition behind eqs (5), (8) and (9) in a diagrammatic analysis and therefore it will be presented using Figures 1 and 2 below.

As discussed earlier, return migration arises because a temporary stay in the host country increases the worker's earning potential in the source country. Therefore migration is a self-selection process which is based on the value of η in this model. The migration flow is composed of negatively selected individuals if $\eta > 1$. In other words, people with lower than average skills in Albania will migrate to EU because in this case only the lower skilled gain the most by moving to the host country. Amongst this cohort of negatively selected individuals, only the more able return to the origin country after a spell in the host country. This case is shown in Figure 1 where we draw the earnings function w^s and w^h (net of migration costs) as thick lines and w^r (net of migration and remigration costs) as dotted line.⁹ Assuming that skills are not perfectly transferable across borders, there are gains from moving for individuals with lower skills, whereas those with relatively higher skills are better off staying in Albania (in terms of eq 8, to satisfy the inequality condition it must be the case that $\mu^h > \mu^s - C^m$). Amongst the lower skilled migrants, only those who have relatively higher skills will face incentives to collect the gains from migration and return to Albania (region A, Fig. 1).

If $\eta < 1$, however, people with skills higher than the average level will migrate. And amongst this pool of positively selected migrants only the relatively less able will find it worthwhile to return after a spell in the host country (region B, Fig. 2).

⁹Where η is the slope of the earnings function in Albania relative to the slope of the earnings function that migrants face in the host country.

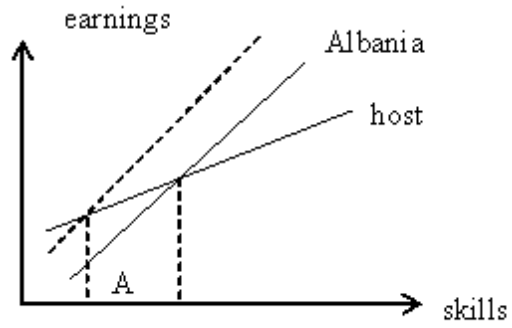


Figure 1: Returns to skills when $\eta > 1$

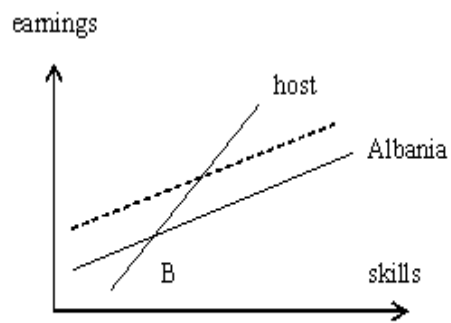


Figure 2: Returns to skills when $\eta < 1$

4 Data and choice of variables

Data used in this paper is based on direct interviews of 1500 individuals in all regions of Albania. The interviews were conducted during the period March 1998-January 1999.¹⁰ Names were randomly selected in the district registers. Numbers attributed by districts are proportional to the size of the district, so the sample is regionally representative. No precise question was asked regarding the location of current residence and therefore it is not clear whether some individuals actually still work abroad but have been interviewed while taking time off in Albania. In order to select only the “real” returnees, we restricted our analysis to those who had migrated and came back at least 2 months before the day of interview.

Moreover, we wanted to avoid the cross-border or seasonal migrants, i.e. those who spend some time of the year abroad and then come back home for the rest of the year. These individuals are defined as persistent migrants and most probably have different characteristics and preferences than the population we want to study (see Constant and Zimmermann (2003) for an investigation of the determinants of repeat migration). Therefore we selected only those individuals who live on earned income, excluding all those who live on remittances (transfers), unemployment benefits, unearned income (i.e. personal savings supposedly earned abroad) or social assistance. We also removed pensioners, housewives and students. Imposing these restrictions of course narrows the scope of our analysis, but it reflects our emphasis on the effect of return migration on the source country labour market.¹¹ Of the original sample of 1500 individuals who were interviewed, selection of valid answers led us to a final sample of 594 wage earners, aged between 16 and 65 (see Table 1 in the Appendix for details on the selection).

Focusing on migrants, we note that less than 30 percent migrated for a period of less than a year, approximately the same percentage migrated for 1 to 2 years, 20 percent for 2 to 3 years, less than 8 percent for 3 to 4 years, 7 percent for 4 to 5 years, while only 7 percent migrated for more than 5 years. Looking at the number of times individuals migrated, we find that 53 percent moved abroad only once, 32 percent did it twice and only 11 percent did it more often. And of those who migrated only once, more than 70 percent did so for more than one year whereas those who migrated twice had an average spell abroad of 13 months each time. The average spell abroad for those who migrated three times is just over ten months. These findings are consistent with the selection of individuals who are return-migrants and not persistent (or seasonal) migrants. Average characteristics are displayed in Table 1.

¹⁰This data was collected within the framework of Phare-ACE project. For a detailed analysis and explanation of the data set see Kule et. al. (2002).

¹¹Individuals out of the labour force are differentiated as pensioners, housewives, students, unemployed. Compared to those included in our paper, they are younger, less often married, less likely to live in cities. Those who have migrated are mainly unemployed and student, those who have no migration experience are mostly students or pensioners.

Table 1: Means of the sample

	Tot. sample	Returned Migrants (204)	Stayers (390)
log(wage)	4.446 (0.644)	4.508 (0.738)	4.413 (0.587)
education	13.973 (2.431)	13.574 (2.363)	14.182 (2.443)
age	37.470 (10.130)	34.843 (9.022)	38.845 (10.414)
male	0.663	0.848	0.567
married	0.714	0.676	0.733
Occupations:			
Managers	0.120	0.113	0.123
Lower man.	0.108	0.088	0.118
Skilled worker	0.222	0.211	0.228
Self-employed	0.207	0.289	0.164
Other paid job	0.253	0.206	0.277
Clerical,unskilled,farmer	0.091	0.093	0.090
Paid in for.currency	0.022	0.044	0.010
Live in cities	0.411	0.466	0.382
Live North	0.146	0.123	0.159
Muslim	0.574	0.574	0.574
Numb. of dependents	0.958 (1.138)	1.123 (1.157)	0.872 (1.120)

The hourly wages converted into US dollars are \$0.72 for the total sample, \$0.81 for return-migrants and \$0.67 for stayers.¹² The Albanian Institute of Statistics (INSTAT) gives the monthly mean income of public sector workers (18% of the labor force) as 10,000 Leks for 1998 while in our sample, using average monthly working hours, the mean average monthly income is 15,351 Leks. We expect this difference to be due to individuals in the private sector earning more than those in the public sector (unfortunately we don't know whether individuals work in the public or private sector in our dataset). The average migrant in our sample is younger, slightly less qualified, less likely to be married and more likely to be male. The differences in average levels of education and age are not statistically significant. Looking at occupation, we note that return-migrants are almost twice as likely to be self-employed as the stayers and there are nearly identical proportion of managers in the stayers and return-migrants sub-populations (12.3% vs 11.3%). These two variables central to our analysis and are therefore discussed in more detail in the empirical section below. Other noticeable difference is the larger proportion of returnees who live in big cities (46% compared to 39%).

¹²The average market rate available for the three quarters of the interview period (II, III, IV, 1998) was 148.8 Lek/\$ (Source: International Financial Statistics, IMF, 2001).

5 Empirical methodology

Two methods are used in order to investigate the issues presented in the theoretical model. We begin by making use of a selection model as proposed by Lee (1978, 1982) and applied to migration by Nakosteen and Zimmer (1980). The model can be summarised by the following three equations:

$$w_i^r = \beta^r x_i + \epsilon_{r_i} \quad (10)$$

$$w_i^s = \beta^s x_i + \epsilon_{s_i} \quad (11)$$

$$m_i^* = \gamma' z_i + u_i \quad (12)$$

The w_i^r is the hourly log wage of individuals who migrated at least once and came back to Albania and w_i^s is the log hourly wage of those who stayed in the country. These hourly wages are explained by a matrix of socio-economic covariates such as education, age and its square, dummy variables for gender, marital status (and its interaction with the gender variable), occupation (managers, lower managers, skilled workers, self-employed, other paid workers and the reference clerical, unskilled and farmers) and a dummy for being paid in a foreign currency.¹³ Equation (12) describes the decision to choose to migrate. The latent variable m_i^* is the difference between benefit and cost from migration (monetary and psychological). Though it is not observed, we know when the individual has decided to migrate, so it can be defined as follows:

$$\text{For migrants} \quad m_i = 1 \quad \text{iff} \quad m_i^* < 0 \quad (13)$$

$$\text{and for non-migrants} \quad m_i = 0 \quad \text{iff} \quad m_i^* \geq 0 \quad (14)$$

Two sets of variables are used to explain the decision to migrate: those included in the wage equations and those not included in them. The second set is needed to identify the model without relying entirely on the normality assumption. To begin with the first one, education is introduced as a variable for the probit migration decision and the wage equation, as this characteristic may be explaining both the migration decision and the wage equation. Age should be negatively associated with the migration decision as older

¹³We have introduced a variable for being paid in foreign currency as we may expect different pay settings for people who work for international organisations or multinational firms than those who work for domestic firms. We observe those who have been abroad at least once are more likely to be hired by private firms (4% of return migrants compared to only 1% of the stayers).

individuals are expected to be more attached to local amenities than younger ones. Furthermore, men are more likely to move than women, a common feature of all studies on migration. The opposite is true for married individuals. We also add an interaction term between gender and marital status as the effect of these variables might be correlated.

As additional variables in the migration equation that are not included in the wage equation, we introduce first the number of dependents within the household with the assumption that tighter liquidity constraints on the household might exert, all else constant, a positive impact on migration decision. The second one is the size of the city where the individual is currently living. Assuming that the individual returned to the place that he/she left when migrating, we expect people living in big cities to be more likely to migrate as family ties might be more relaxed in an urban environment as compared to a rural one. As other identifying variable, we introduce the fact of living in the more mountainous North of the country.¹⁴ Another variable expected to influence migration but not wage is religion. There are two main religions in Albania, Islam and Orthodox Christian. Muslims, who comprise 70% of the population, are expected to face higher (non-pecuniary) costs of migration as opposed to the minority Albanian Orthodox and Roman Catholics (20% and 10% of the population respectively). These costs cover the relatively higher level of difficulty muslims might face in practising their faith in a non-muslim country and also the increased difficulty of assimilation in countries with different religions. We therefore introduced a “muslim” dummy to measure these increased costs of migration for muslims.

The following two conditional wages are defined as the outcome for those who have already made the choice,

$$E(w_i^r | m_i = 1) = \beta^r x_i + E(\epsilon_{ri} | u_i \geq -\gamma' z_i) = \beta^r x_i + \sigma_{e_r} \rho_{ru} \frac{\phi(\gamma' z_i)}{\Phi(\gamma' z_i)} \quad (15)$$

$$E(w_i^s | m_i = 0) = \beta^s x_i + E(\epsilon_{si} | u_i < -\gamma' z_i) = \beta^s x_i + \sigma_{e_s} \rho_{su} \left[-\frac{\phi(\gamma' z_i)}{1 - \Phi(\gamma' z_i)} \right] \quad (16)$$

In order to address the questions posed in the introduction, we need the conditional probabilities for migrants, had they chosen not to migrate and similarly the conditional probabilities of stayers, had they chosen to migrate. Following Maddala (1983), these are given as:

¹⁴As mentioned in Section 2, migration flows originating from the Northern regions of Albania are lower. This can be explained by the increased distance from Greece and/or by more uncertainties surrounding the outcomes on the Greek labour market.

$$E(w_i^s | m_i = 1) = \beta^{r'} x_i + E(\epsilon_{ri} | u_i \geq -\gamma' z_i) = \beta^{r'} x_i + \sigma_{e_r} \rho_{ru} \frac{\phi(\gamma' z_i)}{\Phi(\gamma' z_i)} \quad (17)$$

$$E(w_i^r | m_i = 0) = \beta^{s'} x_i + E(\epsilon_{ri} | u_i < -\gamma' z_i) = \beta^{s'} x_i + \sigma_{e_s} \rho_{su} \left[-\frac{\phi(\gamma' z_i)}{1 - \Phi(\gamma' z_i)} \right] \quad (18)$$

Equation (17) is the conditional wage of stayers, had they chosen to migrate and equation (18) is the conditional wage of migrants, had they chosen to stay. Where $\Phi(\cdot)$ and $\phi(\cdot)$ stand, respectively, for the cumulative and density function of the standard normal, σ_{e_r} and σ_{e_s} are the variances of the error terms of the wage equations for migrants and stayers respectively, and ρ_{su} and ρ_{ru} are the correlations between the stayers and migrants error term, respectively, and that of the migration decision equation. There is no agreement in the literature as to whether these conditional wages should be preferred over the marginal distributions. So in the section devoted to the results we give the marginal effects as well. Average wage differentials can be given for different groups of workers and at different ages and levels of education.

So far we have only been able to give average earning differences whereas the distributional impact of migration might also be of interest to answer the questions posed earlier. One way of identifying the effect of return migration would be to answer the following question: Which density function would prevail if the individual characteristics of migrants had been similar to those of stayers and they had been paid according to the wage schedule observed for stayers? This is one counterfactual density. It is the wage density that would prevail if everybody were receiving stayers' wages. But another way of studying the effect of migration could be to construct a density that would prevail if everybody received migrants' wages. Here the question is: What density would prevail if the characteristics of stayers were similar to those of migrants and they were paid according to the wage schedule of return-migrants? Following DiNardo *et al.* (1996), we can write down these two counterfactuals by the following steps. First we represent the observed density of wages for stayers as the integral of the density of their wages conditional on observed characteristics z over the distribution of these characteristics:

$$g(w|m=0) = \int f^s(w|z)h(z|m=0)dz \quad (19)$$

and similarly for migrants, we have:

$$g(w|m = 1) = \int f^r(w|z)h(z|m = 1)dz \quad (20)$$

We know that the required densities, i.e. the density that would prevail if everybody were receiving stayers wages is:

$$g^s(w) = \int f^s(w|z)h(z)dz \quad (21)$$

and the density that would prevail if everyone were receiving migrants wages is:

$$g^r(w) = \int f^r(w|z)h(z)dz \quad (22)$$

Following Bayes' Law, these densities can be rewritten as¹⁵:

$$g^s(w) = \int \theta^1(z)f^s(w|z)h(z|m = 0)dz \quad (23)$$

$$g^r(w) = \int \theta^2(z)f^r(w|z)h(z|m = 1)dz \quad (24)$$

Note that equations (23) and (24) are similar to equations (19) and (20) except for the weights $\theta^1(z)$ and $\theta^2(z)$ which are respectively:

$$\theta^1(z) = \frac{prob(m = 0)}{prob(m = 0|z)}$$

and

$$\theta^2(z) = \frac{prob(m = 1)}{prob(m = 1|z)} \quad (25)$$

$\theta^1(z)$ can be empirically calculated since $prob(m = 0)$ is simply the proportion of stayers in our sample and $prob(m = 0|z)$ is the probability of being a migrant given individual characteristics which can be estimated by a probit (similar reasoning applies for $\theta^2(z)$). Using these weights, we apply weighted kernel densities to the sample of stayers and migrants to estimate the densities of both counterfactual distributions.

¹⁵The property used is: $h(z) = \frac{h(z|m=0)prob(m=0)}{prob(m=0|z)}$ for the stayers and $h(z) = \frac{h(z|m=1)prob(m=1)}{prob(m=1|z)}$ for the migrants.

6 Results and discussion

6.1 Parametric estimates

Following Ham *et. al.* (2001), we conduct tests on the variables that identify the selection into migrants and stayers. More precisely, we introduce these variables in the wage regressions to check if they are significantly different from zero. If they are significant we exclude them from the entire model, and if they are not significant, we include them in the probit and not the wage estimations. We investigate with four variables: two regional i.e., whether individuals are living in cities and in the North of the country and two personal characteristics: religion (being a muslim) and number of dependents in the household. We expect these variables to affect the migration decision and to be uncorrelated with the error term in the wage equations. We compute Chi-Square tests of their individual and joint significance in the probit and Wald test of their individual and joint significance in the wage equations. The four variables are individually and jointly insignificant in the wage equation for stayers (the individual tests all have a p-value higher than 0.17, and the joint significance is rejected with a p-value of 0.25). For migrants, coefficients for each variable are insignificant (except for living in cities), and test for their joint significance gives a p-value of 0.075 (without the "living in cities" variable, p-value is 0.45). Instruments are jointly significant (p-value of 0 to the second decimal place) in the probit, and they are all significantly different from 0 individually except the "Muslim" variable (p-value of 0.72). The maximum likelihood estimates of the migration model are given in Table 2. For comparison, Tables 2 and 3 in the Appendix also provide estimates of wage equations using Lee (1978) endogenous switching model, with wage equations explained only by education and age, and then adding progressively more exogenous variables. Moreover, in the Appendix, we provide Lee's estimates with only regional characteristics in the probit (first selection rule, Table 4) and then add religion and the number of dependents (second selection rule, Table 5).

6.1.1 Comments on estimates

Note that the estimates for different estimations are rather similar. Generally the coefficients for the stayers' wage equations take the expected sign and are statistically significantly different from zero. One more year of education leads to approximately 4% increase in the hourly wage; age is introduced to measure labor market experience and shows that each subsequent year gives approximately 8.5% increase in the dependent variable. The age profile is concave. One coefficient of interest is the male dummy which is negative and not significant. This result has to be interpreted in the context of an ex-communist country where work was

compulsory for both men and women and wages were set at the national level. Coefficients for occupations take the expected sign with managers earning 66% more than the omitted category (clerical, unskilled and farmers). The premium for self-employment is 52%.

Table 2: Maximum likelihood estimates, second selection rule

Variables	Stayers		Migrants		Migration	
Constant	1.776	(.623)	4.92	(1.003)	.617	(.440)
education	.0405	(.013)	.032	(.028)	-.049	(.025)
age	.0850	(.033)	-.028	(.054)	-.036	(.007)
age squ.	-.094	(.040)	.0595	(.071)		
male	-.0613	(.145)	-.533	(.304)	.979	(.133)
married	-.2636	(.141)	-.110	(.333)	.219	(.145)
male*married	-.0110	(.154)	-.011	(.332)		
Occupations:						
Managers	.6562	(.171)	1.003	(.250)		
Lower man.	.2351	(.174)	.608	(.309)		
Skilled worker	.3588	(.151)	.303	(.238)		
Self-employed	.5171	(.152)	.736	(.221)		
Other paid job	.3007	(.149)	.285	(.230)		
For. currency	-.2696	(.272)	.869	(.180)		
Live in cities					.327	(.115)
Live North					-.345	(.148)
Muslim					-.009	(.113)
Dependents					.0690	(.044)
σ_{e_m}	.875	(.074)				
σ_{e_s}			.5773	(.029)		
ρ_{mu}	-.806	(.087)				
ρ_{su}			-.5678	(.142)		

Notes: The second selection rule is the specification with all four identifying variables mentioned in the text

Interestingly for return-migrants, education and age are not significantly different from zero.¹⁶ However, for migrants, returns to being a manager, self-employed and a “lower” manager are significant and higher than for stayers. Skilled return-migrant workers earn less than skilled stayers. Managers earn between 90 and 100 per cent (depending on the estimation, see Table 2 and Appendix, Tables 2 to 5) more than the omitted category (clerical, unskilled, farmer). The premium for self-employed returners is between 69 and 73 per cent. These results are quite interesting as they suggest that returns to returning take the form of increased earnings in terms of (i) higher positions on the job ladder and (ii) becoming self-employed¹⁷. Better educated and more experienced migrants do not earn higher hourly wages when they return. We observe also a negative and significant sign of the education variable in the migration decision, therefore migration

¹⁶This result is similar to Ham *et. al.* (2001), who found lower return to education for internal migrants in the United States.

¹⁷Overall a Chi-Square test of the joint significance of the occupational variables gives a p-value of zero to the fourth decimal place for stayers and migrants.

is not associated with higher educated individuals. As the theoretical model shows, individuals choose to migrate if the relative rewards to their skills are higher in the host country and then choose to return if they expect the rewards (promotion and/or higher wages etc) to be higher than before in the home country due to newly acquired skills and/or through savings acquired abroad. Therefore returns to skills take the form of access to better jobs in the career ladder but not through returns to formal skills (education and labor market experience). Individuals who choose to migrate and then return face the prospect of access to high paid jobs that do not reward formal training (years of education and labor market experience). The data set shows that 10% of the self-employed and the managers used their savings accumulated abroad to set up a business. This result can therefore be related to the study of Mesnard (2004) who models migration as a way of overcoming constraints of the credit market in the home country. In our context, we observe that individuals who lack formal qualifications required for higher paid jobs tend to migrate to overcome their initial disadvantage. This strategy proves particularly successful as the average earnings of return-migrants are higher than those of stayers.

Looking at the unobserved characteristics, the signs of the corrections for selectivity allow us to draw interesting conclusions. For instance the correction for sample selection in the return-migrant's wage equation is not significant when using a two-step approach (Appendix, Table 4 and 5). The maximum likelihood, however, gives a significant and negative estimate for the correlation coefficient. For stayers, the three estimations give a significant and negative sign for the coefficient of the selectivity variable ($[-\frac{\phi(\gamma'z_i)}{1-\Phi(\gamma'z_i)}]$), which means that the truncation effect is positive. Using the framework of Roy (1951) self-selection model as formalised by Maddala(1983) and others, this indicates that expected earnings of those who choose to migrate may be lower than that of a random individual from the entire sample for given characteristics. And conversely, the expected earnings of those who stayed are higher than the expected earnings of a random individual from the sample. There is positive selection for stayers and support for negative selection of the return-migrants.¹⁸ We expand this issue in the following section where we directly address the question whether the stayers would have performed as well as return-migrants, had they decided to migrate.

6.1.2 Expected earnings and self-selection

Mean income is higher for return-migrants than for stayers by 9 log points, so approximately by 9 percent (see Table 3). Looking at the two counterfactuals, calculated using simple OLS estimations, we note that had they chosen to migrate (and return), stayers' earnings would have been higher than the mean income

¹⁸Again, only those who are currently working.

of return-migrants. The mean earnings of return-migrants, had they chosen to stay, would have been ‘just’ higher than the mean earnings of stayers. However, these estimates are probably biased as they do not take into account the potential self-selection of individuals in either sub-population. Therefore we correct for potential self-selection bias and present the results in columns 2 to 6 of Table 3 which are based on Table 2 and Tables 4 and 5 in the Appendix. For each estimation we give the mean incomes based on the marginal ($E(w^r) = \beta^r x$ and $E(w^s) = \beta^s x$)¹⁹ and the conditional ($E(w^r|m = 1)$ and $E(w^s|m = 0)$) expected wage rate. Marginal distribution should be used for inference on potential migration and conditional distribution should be used for inference on realised migration (Maddala, 1983, p.287). Comparing rows 1 and 2 in Table 3, we observe that migrants made the correct decision in choosing to migrate, as their income is higher than what they would have earned by staying. Comparing the performance return-migrants, had they not migrated, with the performance of stayers (rows 2 and 3), we find that the counterfactual mean income of migrants is always lower than the mean income of stayers. This shows that the performance of return-migrants, if they had stayed, would have been worse than that of the stayers. As for the stayers, comparing rows 3 and 4 it can be seen that their mean income would have been higher had they spent time abroad. The order of this advantage is .17 and 1.16 log points using the marginal and conditional expected means respectively of the Maximum likelihood estimation. In the framework of our theoretical model, we observe that more skilled individuals do not migrate if their potential earnings net of migration costs are lower in the host countries compared to their wage at home. Our results give rise to a story of the more able/skilled individuals in Albania facing higher assimilation costs in the host labour markets. This may come from the difficulty to practice their profession in a foreign language. This, for instance, would apply to such professions as medical doctors, lawyers or teachers. For the less skilled, such costs may be much lower as the jobs performed in the host countries do not require a high fluency of the foreign language.²⁰ These results show that return-migrants are negatively selected as depicted in the theoretical analysis in fig 1.

¹⁹We choose the term "marginal" rather than "unconditional" following Maddala(1983) and Van der Gaag and Vijverberg (1991).

²⁰Our data set offer some help in identifying this increased assimilation costs faced by the stayers. They are asked to give the main reason why they did not migrate amongst eight possible answers. The results are as follows: "family" (16%), fear of losing the current job (12%), not having a visa (11%), love for the home country (9%), only at the sixth place comes the financial cost (6%), then being too old (5%) and health reasons (2%). No one chose the risk of losing social assistance.

Table 3: Estimated mean hourly wage for return migrants and stayers

	OLS	Endogenous switching, Lee(1978)		Maximum likelihood			
		First sel. rule marg.	Sec. sel. rule cond.	Sec.sel rule marg.	Sec.sel rule cond.		
Migrants (204 cases)							
Mean income	4.5	4.96	(4.5)	4.73	(4.5)	5.15	4.5
Had they stayed (counterfact.)	4.42	3.91	3.21	4.02	3.47	4.20	3.90
Stayers (390 cases)							
Mean income	4.41	4.04	4.41	4.13	4.41	4.26	4.41
Had they migrated (counterfact.)	4.49	5.10	(5.34)	4.80	(4.91)	5.32	5.66

Notes: Marg. is for: $E(w^r) = \beta^r x$ and $E(w^s) = \beta^s x$
 and Cond. is for: $E(w^r|m = 1)$ and $E(w^s|m = 0)$

6.2 Results using semi-parametric estimates

We now investigate the entire density of hourly wages. All graphs presented here give estimates calculated with a Gaussian kernel function. We used the Silverman (1986, eq. 3.31) procedure to select the optimal bandwidth, its value lies at around 0.14. Kernel estimates for the entire sample, for the stayers and for the return-migrants, are displayed in Figure 3. In Figure 4, densities for the total sample are decomposed into the weighted sum of the densities of return-migrants and stayers. We simply multiply the sub-group densities of Figure 3 by the sub-group population shares.

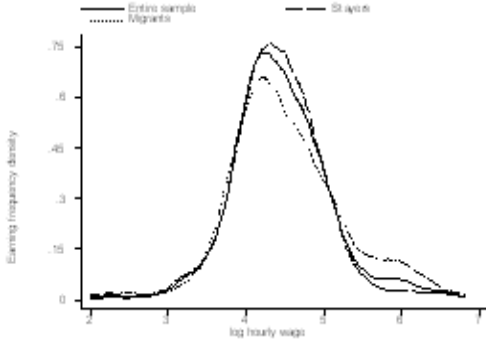


Figure 3: Kernel densities

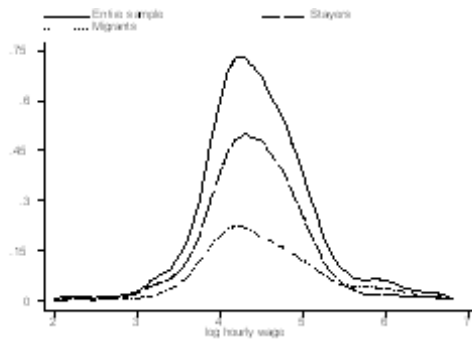


Figure 4: Weighted densities

Figure 3 shows that return-migrants tend to account for a larger part of the total distribution at higher hourly wages. There is clearly a clustering of the distribution at higher wages for those who have migrated

and leads to a small "bump" at the top of the overall distribution. These observations based on the raw distributions are interesting but cannot reveal the real effect of migration as we compare subpopulations with rather different characteristics. We already know from Table 1 that migrants tend to be less educated, younger and, more often, male.

The different curves may be due more to these individual characteristics than to migration. So we have to go a step further in comparing populations with similar characteristics. This could be done in two ways, either by displaying the distribution of wages as if everyone were paid the stayers' wage, or by graphing the distribution of wages as if everyone were paid the return-migrants' wage. More precisely, in the first case we answer the following question: Which density function would prevail if individual characteristics of return-migrants had been similar to those of stayers and they had been paid according to the wage schedule observed for stayers? This is done in Figure 5 which gives the hypothetical counterfactual density together with the density of the entire population. The difference between the two curves can be interpreted as the effect of return migration. The curve called the density without migration is calculated using eq (24). Figures 1 and 4 in the Appendix present the propensity scores of the probit and also the weights $\theta^1(z)$ and $\theta^2(z)$. Note that the counterfactual density in Figure 5 is rather similar to the density of the entire sample. Had the return-migrants been paid the same as the stayers and their characteristics would have been similar, we would have observed a slightly different density function. Mainly the small cluster at the top of the distribution disappears and is compensated by a shift of the curve to the right just after the mode of the distribution. So interpreting the effect of return migration as the difference between the two curves, we can say that its effect is rather reduced at the bottom of the distribution and can explain the bump at around 6 log hourly Lek.

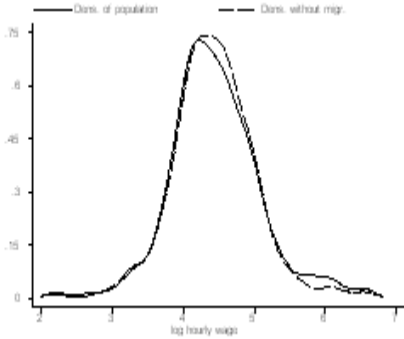


Figure 5: Hypoth. density without migration

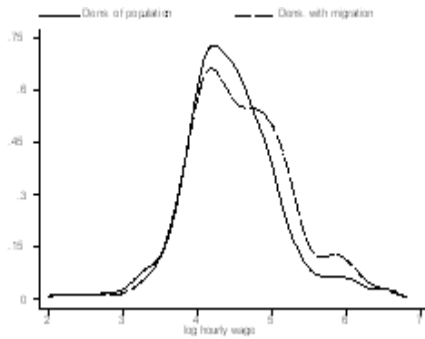


Figure 6: Hypoth. density with migration

Figure 6 gives complementary information as here the reference is the return-migrants sub-population. The counterfactual curve is now the density that would have prevailed if the characteristics of return-migrants were similar to those of stayers. This would have resulted in the density function lying to the right of the actual one. This counterfactual distribution is nearly bimodal, with a second (lower) mode at higher wage. These figures give more support to the negative selection of return migrants. In particular, we observe here that the effect of migration would have been much stronger had the return-migrants characteristics been more similar to those of stayers.

6.3 Results with Disaggregated Characteristics

In this section we want to check that the above results, which are based on the mean income of all individuals, still hold if the individuals are disaggregated by qualification levels, age and type of employment (self- or wage employment). Using the maximum likelihood estimates, we therefore calculate the marginal and conditional expected hourly wages for three different characteristics: Those with more and less than 14 years of schooling, those more and less than 30 years of age, and for wage and self-employed workers (see Appendix, Table 6).²¹

The first cell of first column of Table 4 shows that the stayers, had they migrated, would have earned 1.17 log points more than the return-migrants' actual earnings. And the first cell of column 3 shows that the return-migrants, had they decided to stay, would have earned .42 points less than the actual earnings of stayers. These results strongly suggest that the sub-population of stayers is composed of better performers. For all decomposition of the population, by age, employment and level of education, stayers would have performed better had they migrated. We observe that highly educated (young and old) stayers would have gained more, had they decided to move than low educated and compared to similar migrants. Also highly educated return-migrants (young and old) would have lost more, had they stayed, compared to stayers with same education level.

Another area of interest is to look at the individual comparative advantage for each sub-population. Here, comparison is made between what the individuals would have earned (had they decided otherwise) with what they are actually earning. So the first cell of first column of Table 5 shows that low educated stayers are earning 1.24 log points less than what they would be earning, had they decided to move. And the first cell of column 3 implies that the less qualified return-migrants earn .59 than if they had chosen not to migrate. The results confirm that for each type of characteristics migrants made the right decision. However,

²¹Dustmann and Kirchkamp (2002) showed that return migrants choosing between self-employment or wage sector tend to experience different outcomes when they return.

Table 4: Absolute advantage for different characteristics

	Stayers		Migrants	
	col.1	col.2	col.3	col.4
	low educ.	high educ.	low educ.	high educ.
wage	1.17	1.26	-.42	-.54
self	1.12	1.12	-.50	-.54
young	1.08	1.46	-.29	-.45
old	1.04	1.34	-.42	-.57

Notes: Absolute advantage is computed as the difference between log hourly wages of stayers (migrants) and the counterfactual mean earnings of migrants (stayers).

Table 5: Comparative advantage for different characteristics

	Stayers		Migrants	
	col.1	col.2	col.3	col.4
	low educ.	high educ.	low educ.	high educ.
wage	-1.24	-1.37	.59	.65
self	-1.21	-1.09	.69	.51
young	-1.32	-1.50	.53	.73
old	-1.19	-1.35	.57	.58

Notes: Comparative advantage are computed as the difference between mean log hourly wages and the counterfactual for each population.

as mentioned in the previous section, the stayers must face unobserved costs of migration, which prevent them to migrate despite the fact that they would have been financially better off in doing so. Therefore the results found earlier on the aggregated sub-population (Table 3) are not affected when we take into account the different characteristics.

7 Conclusion

Data on Albanian migration suggests that a sizeable proportion of migrants return to Albania after a short spell abroad. This predominantly return behaviour of Albanian migration offers an interesting case study to investigate the effect of migration on the source country labour market. Using a sample of 694 individuals active on the Albanian labour market, we compare those who returned after a spell abroad (204 individuals) with those who never migrated (390 individuals).

We have investigated the negative or positive selection of return-migrants by comparing their performance in the source country with those of the stayers in the framework of the Roy theoretical model of self-selection. We found support for the negative selection of return-migrants. Using counterfactual analysis, we found that

had the stayers decided to migrate and return, they would have earned a higher hourly wage (in the order of .17 to 1.09 log points) than the return-migrants. Applying the semi-parametric approach of DiNardo et. al. (1996), it was shown that return migration results in a slight rightward shift of the wage distribution in Albania. However, had the characteristics of migrants been more similar to those of stayers, the wage distribution would have markedly moved to the right. We interpret this result as further supporting evidence for the negative selection of return-migrants compared to the stayers. We explain the choice of stayers by their higher costs of migration. Being on average more skilled, they would face higher assimilation costs in the host countries such as knowledge of the host country language and the recognition of their formal training acquired at home. For typical low skilled migrants such costs are much lower as they are expected to be active in menial jobs where few contacts and relatively little training are required. We also observed that rewards to the typical human capital variables, age and education, are not statistically significant in the home country labour market for the return-migrants whereas the opposite prevails for the stayers.

This paper adds to the scant literature on the self-selection process characterising the flows of return migrants in the context of the source country labour market. Albania is a relatively poor and small country with a dominant agricultural sector typical of a large number of Central and Eastern European and developing countries. We may expect our results to apply to similar countries as well.

As potential policy implications, we may mention the increased hourly wage of returnees due to their spell abroad, despite them appearing to be negatively selected. This is clearly beneficial for the source country economy, especially as a large proportion of the returnees appear to choose to set up successfully as self-employed. It could therefore be inferred from this behaviour that credit constraints play an important role in the decision to leave, work and save abroad, and then return to participate in the local economy. It seems, therefore, that better access to credit market will be helpful in promoting higher pay-off to self-employment in Albania.

For the host countries, a common worry has been the fear of the adverse effect of large flows of unskilled immigrants entering their labour markets. It appears that, at least in the case of Albania, a large proportion of immigrants choose the short-term (or guest worker) option. Nevertheless, it may be advisable for countries fearing these adverse effects to implement short-term work permits to be able to better monitor such flows. Finally, host countries could as well try to lessen the incoming flows by favouring the creation of micro-credit institutions in the source countries.

8 References

- Albanian Center for Economic Research (2002) *Common Country Assessment: Albania*. Prepared for the United Nations System in Tirana, Albania. url: <http://www.undp.org.al/download/pdf/cca2002.pdf>
- Bauer T, Pereira P, Vogler M, Zimmermann K (2002) Portuguese Migrants in the German Labor Market: Performance and Self-Selection. *International Migration Review*, 36(2): 467-491.
- Borjas G (1987) Self-Selection and the Earnings of Immigrants. *American Economic Review* 77(4): 531-553.
- Borjas G (1991) Immigration and Self-Selection. In: Abowd J, Freeman R (eds) *Immigration, Trade and the Labor Market*. Chicago University Press, Chicago, 29-76.
- Borjas G (1999) Immigration and Welfare Magnets. *Journal of Labour Economics* 17(4): 607-637.
- Borjas G, Bratsberg B (1996) Who Leaves? The Out-Migration of the Foreign-Born. *The Review of Economics and Statistics* 78(1): 165-176.
- Chiswick B (1999) Are Immigrants Favorably Selected?. *American Economic Review* 89(2): 181-5.
- Co C, Gang I, Yun M-S (2000) Returns to Returning. *Journal of Population Economics* 13(1): 57-79.
- Constant A, Massey D (2003) Self-selection, Earnings and Out-Migration: A Longitudinal Study of Immigrants to Germany. *Journal of Population Economics* 16(4): 631-653.
- Constant A, Zimmermann K (2003) The Dynamics of Repeat Migration: A Markov Chain Analysis. *Discussion Paper No. 885*, IZA, Bonn.
- DiNardo J, Fortin N, Lemieux T (1996) Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semi-Parametric Approach. *Econometrica* 64(5): 1001-1044.
- Dustmann C (2003) Return Migration, Wage Differentials and the Optimal Migration Duration. *European Economic Review* 47(2): 353-369.
- Dustmann C, Kirchkamp O (2002) The Optimal Migration Duration and Activity choice after Re-migration. *Journal of Development Economics* 67(2): 351-372
- Jarvis C (2000) The Rise and Fall of the Pyramid Schemes in Albania. *IMF staff papers* 47(1), Washington, D.C.
- Ham J, Li X, Reagan P (2001) Matching and Selection Estimates of the Effect of Migration on Wages for Young Men. Mimeo, Ohio State University.

- Harris J, Todaro M (1970) Migration, Unemployment and Development: A Two Sector Analysis. *American Economic Review* 60(1): 126-142.
- Hartog J, Winkelmann R (2003) Comparing Migrants to Non-Migrants: The Case of Dutch Migration to New Zealand. *Journal of Population Economics* 16(4): 683-705.
- Heckman J (1979) Sample Selection as a Specification Error. *Econometrica* 47(1): 153-161.
- Heckman J, Honore B (1990) The Empirical Content of the Roy Model. *Econometrica* 58(5): 1121-1150
- International Organisation for Migration (1995) Profiles and Motives of Potential Migrants from Albania. Geneva.
- Kule D, Mançellari A, Papapanagos H, Qirici S, Sanfey P (2002) The Causes and Consequences of Albanian Emigration during Transition: Evidence from Micro Data. *International Migration Review* 36(1): 229-239.
- Lee LF (1978) Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables. *International Economic Review* 19(2): 415-433
- Lee, LF (1982) Some Approaches to the Correction of Selectivity Bias, *Review of Economic Studies*, 49(3): 355-372.
- Leon-Ledesma M, Piracha M (2001) International Migration and the Role of Remittances in Eastern Europe. *University of Kent Discussion Paper No.13/01*.
- Maddala G (1983) *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge University Press, Cambridge.
- Mesnard A (2004) Temporary Migration and Capital Market Imperfections. *Oxford Economic Papers* 56(2): 242-262.
- Nakosteen R, Zimmer M (1980) Migration and Income: The Question of Self-Selection. *Southern Economic Journal* 46(3): 840-51.
- OECD (2000) *Trends in International Migration*, SOPEMI, Paris.
- Papapanagos H, Sanfey P (2001) Intention to Emigrate in Transition Countries: The Case of Albania. *Journal of Population Economics* 14(3): 491-504
- Roy A (1951) Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers* 3(2): 135-146.
- Silverman B (1986) *Density Estimation in Statistics and Data Analysis*. Chapman and Hall, London.
- Treichel V (2002) Stabilization Policies and Structural Reform in Albania Since 1997. *IMF Policy Dis-*

Discussion Paper 02/02, Washington, D.C.

Van der Gaag J, Vijverberg W (1988) A Switching Regression Model for Wage Determinants in the Public and Private Sectors of a Developing Country. *Review of Economics and Statistics*, 70(2): 244-252

Appendix

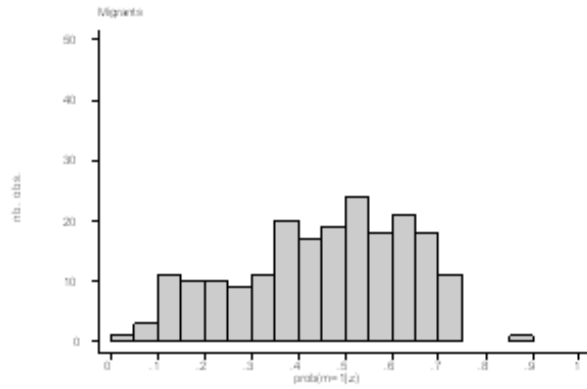


Figure A1: Estimated prop. scores

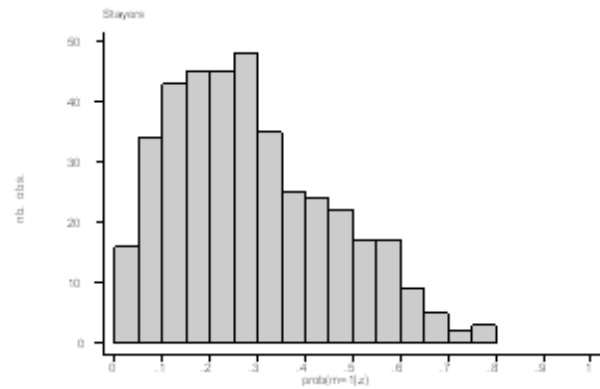


Figure A2: Estimated prop. scores

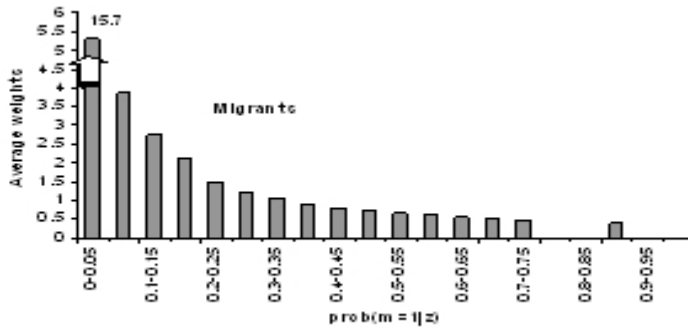


Figure A3: Est. prop.scores and weights ($\theta^2(z)$)

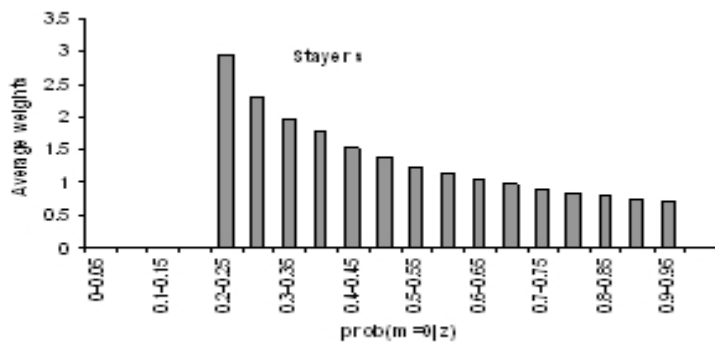


Figure A4: Est. prop. scores and weights ($\theta^1(z)$)

Table 1: Sample Selection Criteria

1500	Total numbers of interviews
-37	Not reporting their age
-25	Individuals less than 15 or more than 65 years old
-460	Retired, not active, student, unemployed, missing occupation
-33	Not reporting years of education
-186	Not reporting earned income
-71	Missing or non valid working hours
- 88	Migrants returned since less than two months
- 6	Hourly wage higher than the 99th percentile
594	204 return migrants and 390 stayers

Table 2: Migrants, different specification of the wage function, dep. variable: Lw, second selection rule

Variables	Coeff.	St.-Err.	Coeff.	St.-Err.	Coeff.	St.-Err.	Coeff.	St.-Err.
Constant	4.224	(0.888)	4.455	(0.945)	4.477	(0.969)	4.536	(0.912)
education	0.045	(0.023)	0.049	(0.024)	0.050	(0.024)	0.025	(0.026)
age	-0.018	(0.047)	-0.018	(0.047)	-0.006	(0.050)	-0.030	(0.047)
age squ./100	0.029	(0.068)	0.034	(0.063)	0.023	(0.065)	0.050	(0.061)
male			-0.183	(0.251)	-0.380	(0.372)	-0.234	(0.345)
married					-0.241	(0.312)	-0.038	(0.300)
male*marr.					0.180	(0.319)	-0.014	(0.304)
Managers							0.900	(0.235)
Lower Man							0.571	(0.238)
Skilled work							0.299	(0.198)
Self-emp.							0.691	(0.188)
Other job							0.261	(0.199)
For.curr.							0.686	(0.239)
Lambda	-0.099	(0.182)	-0.286	(0.316)	-0.377	(0.344)	-0.251	(0.319)

Table 3: Stayers, different specification of the wage function, dep. variable: Lw, second selection rule

Variables	Coeff.	St.-err.	Coeff.	St.-err.	Coeff.	St.-err.	Coeff.	St.-err.
Constant	2.106	(0.486)	2.129	(0.512)	1.371	(0.576)	1.487	(0.560)
education	0.055	(0.012)	0.054	(0.013)	0.055	(0.014)	0.045	(0.015)
age	0.067	(0.023)	0.067	(0.023)	0.104	(0.026)	0.089	(0.025)
age squ./100	-0.078	(0.029)	-0.079	(0.030)	-0.112	(0.031)	-0.029	(0.031)
male			0.017	(0.124)	-0.057	(0.169)	-0.170	(0.166)
married					-0.243	(0.122)	-0.293	(0.120)
male*marr.					-0.058	(0.131)	-0.031	(0.126)
Managers							0.676	(0.129)
Lower Man							0.237	(0.127)
Skilled work							0.377	(0.109)
Self-emp.							0.548	(0.116)
Other job							0.322	(0.112)
For.curr.							-0.259	(0.267)
Lambda	-0.329	0.124	-0.298	0.246	-0.543	0.271	-0.599	(0.265)

Table 4: Endogenous switching model, Lee(1978), first selection rule

Variables	Stayers		Migrants		Migration	
Constant	1.281	(0.625)	4.734	(0.936)	0.641	(0.393)
education	0.047	(0.017)	0.033	(0.027)	-0.045	(0.024)
age	0.094	(0.027)	-0.028	(0.047)	-0.037	(0.007)
age squar.	-0.001	(0.000)	0.001	(0.001)		
male	-0.239	(0.193)	-0.437	(0.395)	1.003	(0.132)
married	-0.317	(0.133)	-0.106	(0.302)	0.253	(0.148)
male*married	-0.029	(0.133)	0.027	(0.298)		
Occupations:						
Managers	0.669	(0.136)	0.902	(0.233)		
Lower man.	0.240	(0.134)	0.562	(0.236)		
Skilled worker	0.367	(0.114)	0.301	(0.197)		
Self-employed	0.545	(0.121)	0.693	(0.187)		
Other paid job	0.310	(0.118)	0.254	(0.197)		
For. currency	-0.230	(0.282)	0.681	(0.236)		
Live in cities					0.247	(0.119)
Live North					-0.365	(0.171)
lambda	-0.758	(0.326)	-0.497	(0.397)		

Table 5: Endogenous switching model, Lee(1978), second selection rule

Variables	Stayers		Migrants		Migration	
Constant	1.487	(0.561)	4.536	(0.912)	0.717	(0.413)
education	0.045	(0.015)	0.025	(0.026)	-0.049	(0.024)
age	0.089	(0.025)	-0.030	(0.047)	-0.037	(0.007)
age squ.	-0.001	(0.000)	0.000	(0.001)		
male	-0.170	(0.167)	-0.234	(0.345)	0.975	(0.133)
married	-0.293	(0.120)	-0.038	(0.300)	0.219	(0.150)
male*married	-0.031	(0.127)	-0.014	(0.304)		
Occupations:						
Managers	0.676	(0.130)	0.900	(0.235)		
Lower man.	0.237	(0.128)	0.571	(0.238)		
Skilled worker	0.377	(0.109)	0.299	(0.198)		
Self-employed	0.548	(0.117)	0.691	(0.188)		
Other paid job	0.322	(0.113)	0.261	(0.199)		
For. currency	-0.259	(0.267)	0.686	(0.239)		
Live in cities					0.267	(0.121)
Live North					-0.421	(0.176)
Muslim					-0.126	(0.117)
Dependents					0.095	(0.051)
lambda	-0.599	(0.265)	-0.251	(0.319)		

Table 6: Mean and Counterfactual mean incomes for different characteristics

Migrants (204 cases)				Mean		Counterfactual	
	Educ	Age	Wage	Marg	Cond.	Marg	Cond.
				col.1	col.2	col.3	col.4
row 1	<14	all	1	5.02	4.45	4.13	3.86
2	>14	all	1	5.42	4.66	4.36	4.01
3	<14	all	0	4.93	4.36	4.03	3.77
4	>14	all	0	5.14	4.40	4.23	3.89
5	<14	<30	1	4.86	4.40	3.98	3.77
6	>14	<30	1	5.38	4.74	4.30	4.01
7	<14	>30	1	5.15	4.50	4.24	3.93
8	>14	>30	1	5.45	4.60	4.41	4.02
Stayers (390 cases)				Mean		Counterfactual	
	Educ	Age	Wage	Marg	Cond.	Marg	Cond.
row 9	<14	all	1	4.10	4.28	5.12	5.52
10	>14	all	1	4.40	4.55	5.60	5.92
11	<14	all	0	4.11	4.27	5.12	5.48
12	>14	all	0	4.31	4.43	5.25	5.52
13	<14	<30	1	3.93	4.16	4.99	5.48
14	>14	<30	1	4.28	4.46	5.49	5.86
15	<14	>30	1	4.18	4.35	5.18	5.54
16	>14	>30	1	4.45	4.59	5.64	5.94

Table 7: Absolute advantage for different characteristics, marginal

	Stayers		Migrants	
	col.1	col.2	col.3	col.4
	low educ.	high educ.	low educ.	high educ.
wage	.10	.18	.03	-.04
self	.19	.11	-.08	-.08
young	.13	.11	.05	.02
old	.03	.19	-.06	-.04

Notes: Absolute advantage is computed as the difference between log hourly wages of stayers (migrants) and the counterfactual mean earnings of migrants (stayers).

Table 8: Comparative advantage for different characteristics, marginal

	Stayers		Migrants	
	col.1	col.2	col.3	col.4
	low educ.	high educ.	low educ.	high educ.
wage	-1.02	-1.20	.89	1.16
self	-1.01	-.94	.90	.91
young	-1.06	-1.21	.88	1.08
old	-1.00	-1.19	.91	1.04

Notes: Comparative advantage are computed as the difference between mean log hourly wages and the counterfactual for each population.