EMIGRATION AND HUMAN CAPITAL: WHO LEAVES, WHO COMES BACK AND WHAT DIFFERENCE DOES IT MAKE?

2006

Aitor Lacuesta

Documentos de Trabajo
N.º 0620

BANCO DE ESPAÑA

EMIGRATION AND HUMAN CAPITAL: WHO LEAVES, WHO COMES BACK AND WHAT DIFFERENCE DOES IT MAKE?

EMIGRATION AND HUMAN CAPITAL: WHO LEAVES, WHO COMES BACK AND WHAT DIFFERENCE DOES IT MAKE? (*)

Aitor Lacuesta (**)

BANCO DE ESPAÑA

(*) This paper is a revised part of my dissertation at the University of Chicago. I thank Robert LaLonde, Steve Levitt and especially Casey Mulligan for their valuable advice. I would also like to thank Gary Becker, David Cuberes, Christina Gathmann, Angel Gavilan, Ricard Gil, John Guryan, Salvador Navarro, Josep Vilarrubia and Gema Zamarro for helpful comments and discussions. At the same time, I would like to thank Christian Borja and Salvador Navarro, who kindly directed my attention to the Mexican data. The Fundación Rafael del Pino and the Fundación Ramón Areces are acknowledged for providing financial aid during my years at Chicago. The opinions and views expressed in this paper do not necessarily reflect those of the Banco de España.

(**) Aitor Lacuesta. DG Economics, Statistics and Research, Banco de España. C/ Alcalá, n.º 48. 28014 Madrid, Spain. aitor.lacuesta@bde.es.

Documentos de Trabajo. N.º 0620
2006

ISSN: 0213-2710 (print)
ISSN: 1579-8666 (on line)
Depósito legal: M.36564-2006
Imprenta del Banco de España

Abstract

This paper studies the loss of human capital that emigration generates in the country of origin. To that end I estimate the human capital distribution of emigrants had they not migrated. Unlike previous studies, I take into account the selection of migrants in terms of unobserved characteristics that affect their productivity. Wages in Mexico of those migrants who come back home after being abroad for some time will be crucial to learn something about the selection of non-returning migrants in terms of unobserved productivity. To test whether returning migrants' wages contain any useful information, I follow two steps. First, I use the model of Borjas and Bratsberg (1986) to show that, regardless of the cause for coming back, the distribution of abilities of non-returning migrants is more similar to the distribution of temporary migrants than to that of non-migrants. Moreover, I test some implications of the model in the data. Second, I show that returning migrants' wages reflect their pre-emigration productivity and are not affected by possible human capital gains derived from the decision to emigrate. Taking into account all this evidence, I use returning migrants' wages in Mexico upon return to estimate the distribution of human capital of non-returning migrants had they not migrated. I show that emigrants come form the middle part of the distribution of human capital in the origin country. I find evidence that taking unobserved human capital factors into account is relevant for the dispersion of the estimated distribution as well as for each of its quantiles. Moreover, it does not greatly affect the aggregate mean of human capital.

JEL Classification: C14, J10, J31.

1 Introduction

This paper describes the loss of human capital that emigration generates in the country of origin. The main difference from other papers analysing the same question is that I take into account the selection of migrants in terms of courage, adaptability and other unobserved variables which might be correlated with productivity. Additionally, unlike most previous economic analyses, I consider migration as a reversible action. Returning migration is going to be crucial to analyse the selection in terms of unobserved factors of non-returning migrants.

When analysing the overall effect of emigration on the origin labour market it is important to realise that this does not only depend on the number of individuals that are lost but also on their characteristics. If migrants come from the upper part of the distribution of aggregate human capital, the effect on the aggregate level of human capital of this movement is much more negative than if they come from the lower part of it. This is why, from a policy perspective, it is very important to obtain a good estimation of their productivity at home.

Census data provides information on the self-selection of migrants regarding age and education, which are very important variables for productivity. Evidently, if these two variables fully describe the individual's productivity at home, and assuming that the labour market is competitive[1], we can estimate the loss of human capital that emigration generates in the country of origin using the wage at home of a non-migrant. This is the strategy of Hanson and Chiquiar (2005), who use a method presented in DiNardo et al. (1996) to estimate the wage distribution of migrants in the country of origin $f_{mig}(w^{origin} \mid x)$. Their counterfactual estimation $g_{mig}(w^{Origin} \mid x)$ is defined as the wage that a person with the same characteristics as a migrant would earn if he earns as much as a non-migrant. Hanson and Chiquiar (2005) show that the counterfactual distribution could be written as the wage density of non-migrants re-weighted by a factor θ:

$$g_{mig}(w^{Origin}) = \int \theta f_{non\_mig}(w^{Origin} \mid x) h_{non\_mig}(x) dx$$

where θ is a re-weighting function that expresses the different mix of characteristics for each particular group.

$$\theta = \frac{\Pr(i = mig \mid x)}{\Pr(i = non\_mig \mid x)}$$

However, age and education do not fully describe the productivity of migrants at home. There are many more variables that are not observed by the econometrician and that might affect someone's productivity such as innate ability, language skills, informal training, courage, adaptability and self-esteem. Each of them affects in some way the conditional

1. This is necessary in order to equalise productivity and wages.

distribution of wages, which makes the estimation of the loss of human capital derived from emigration much more difficult, since these variables might be also correlated with the decision to leave. Actually, we would expect emigrants to be more able to adapt to a different environment, to have better language skills and to be more courageous than non-migrants. Consequently, the previous estimation is not correct because the conditional density of wages is affected by the migration status of the person[2].

$$f_{mig}(w^{Origin} \mid x) \neq f_{non\_mig}(w^{Origin} \mid x)$$

A priori the omission of unobserved ability produces an unclear bias in the estimation of the mean human capital of migrants because there is no consensus in the literature about whether migrants are positively or negatively selected in this dimension[3]. Moreover, the bias affects the whole distribution of abilities, which means that other moments and quintiles of the distribution that might be relevant in answering policy-related questions might also be wrongly captured.

In order to take this issue into account, information about wages in the country of origin of non-returning migrants prior to leaving would be ideal. However, census data only provides information about the wage in the actual place of residence and there is nothing said about the previous salaries at home. Another strategy consists of searching for a set of individuals in the source country who were equal in terms of productivity to permanent migrants. I would require an instrument, i.e. a variable unrelated to productivity that makes people stay in the country of origin when otherwise they would have migrated. However, with census data there is not much information about individuals that may be used as a proper instrument.

Fortunately, many censuses provide the wage at home of returning migrants. A priori, returning migrants might be more similar to non-returning migrants than to non-migrants since they also decided to leave their country of birth at some point. Nevertheless, they are not a perfect control group for two reasons: first, they decided to go back home after a while instead of staying in the host country, which means that this subset of the population is again selected somehow relative to non-returning migrants. Second, the productivity of temporary migrants might have been affected by their living abroad for a period of time. If this was the case, their wage at home upon return $f_{ret\_mig}(w^{origin} \mid x)$ would be a poor proxy for the wage of non-returning migrants prior to migration.

Therefore, in order to test the validity of using returning migrants' wages to estimate the human capital of non-returning migrants, I need to verify 3 issues:

1) First, I need to be sure that, theoretically, the selection of non-returning migrants is more similar to the selection of returning migrants than to that of non-migrants. In order to do that, I use the model of Borjas and Bratsberg (1996), who analyse the selection of returning, non-returning and non-migrants in terms of unobserved ability. The previous model provides for two motives for returning: productivity gains and non-realised expectations. Regardless of

2. DiNardo et al. (1996) already made this point when presenting the technique.
3. Borjas (1987) predicts a negative selection, whereas Chiswick (1978) predicts a positive one.

the reason underlying the cause of return, returning migrants are placed between non-migrants and non-returning migrants in terms of unobserved ability.

2) Second, I need to check that the model holds in the data. Since Mexican migration to the United States has been quantitatively significant and has presented big returning rates compared to other origins, the empirical part has focused on this particular case. I add some observable characteristics to the model in order to obtain testable implications. In particular, I show that the probability of migration to the United States and the probability of coming back to Mexico given that the person migrated vary with characteristics in opposite directions, as predicted in the model.

3) Finally, I need to ensure that there are no human capital gains derived from emigration. Using data from some rural areas in Mexico with big outflows to the United States, I show that migrants do not use migration as a means of changing the sector they were working in before emigration, at least not in the short run. Moreover, analysing how the wage gap between non-migrants and returning migrants varies with the duration of the migration, I show that human capital gains, if any, do not appear to be very significant compared to pre-migration differences in terms of unobserved ability.

All this evidence reinforces the idea that we may estimate wages of non-returning migrants using wages of temporary migrants in order to partially incorporate the unobserved heterogeneity that Hanson and Chiquiar were lacking. The new estimation of the migrants' human capital distribution turns out to be much more concentrated than that of Hanson and Chiquiar. Instead, the mean human capital does not change much in relation to the mean component of Hanson and Chiquiar's estimation. In both cases, when we compare migrants with non-migrants, the first group is more concentrated in human capital levels ranging in the middle of the distribution.

The structure of the paper is as follows: in the second section I set out the model of Borjas and Bratsberg (1986) with their main implications for selection. In the third section, I test some implications of the model in the data. The fourth section analyses whether there is a change in human capital derived from emigration and, once this possibility is rejected, the fifth section presents the estimation of the non-returning migrants' distribution of human capital. The sixth and final section concludes.

2 Model

In this section I set out the model in Borjas and Bratsberg (1996) to analyse the selection of returning, non-returning and non-migrants in terms of unobserved ability. Individuals migrate to maximise their income. They might migrate with the idea of doing so temporarily to increase their productivity abroad and go back home afterwards to get the returns, or they might decide to leave forever. Once abroad, if things are not as expected, they might decide to change their initial plans (either to stay longer than expected or to come back immediately without the human capital gain).

Everybody has a stock of human capital that is composed by observed factors *X* such as gender, age or education and unobserved factors that are gathered under the letter *v*. The returns to each factor are different depending on the location of the individual. If i denotes the place of residence, $h^i(X,v)$ will be the payoff for a given set of characteristics. Since the empirical part refers to Mexico and the United States, the superscript *Mex* and *US* refer to those two places respectively. The function $h^i(X,v)$ fully characterises the wage at home[4], but the wage abroad cannot be estimated with certainty. This is justified because the information on job opportunities abroad is not perfect. Therefore, the actual wage is not known until the person goes to the United States when a shock that affect someone's wage ε is realised. The functional forms for the wages are:

$$w^{Mex} = \mu^{Mex}(X) + v$$

$$w^{US} = \mu^{US}(X) + \rho v + \varepsilon$$

The person might also decide to stay abroad for a while. The experience abroad increases his productivity by *k*. In that case, the life cycle earnings assuming that the person needs a fraction of time π in the US to increase his productivity are defined as:

$$w^{US-Mex} = \pi w^{US} + (1-\pi)(w^{Mex} + k)$$

The individual must choose his location in order to maximise his income net of migrating costs. Let M be the "time equivalent" measure of the costs of migrating and R the "time equivalent" cost of returning. Hence, assuming risk neutrality a person migrates to the United States if:

$$\max\left[Ew^{US} - M, Ew^{US-Mex} - M - R\right] > w^{Mex}$$

4. This is a plausible assumption since most migrants leave the country after having worked for a while.

and a person migrates and then comes back if:

$$\max\left[Ew^{US} - M, Ew^{US-Mex} - M - R\right] > w^{Mex} \text{ and}$$

$$\max\left[w^{Mex} - R, Ew^{US-Mex} - R\right] > w^{US}$$

Notice that there is one equation that coincides for temporary and permanent migrants. On the other hand, temporary migrants have another restriction that comes from the two possible motivations to come back (prior expectation of coming back, actual realisation of expectations). Notice that if the person decides to go back home immediately, there is no human capital gain derived from emigration. With this simple model the authors obtain a sorting in terms of unobserved abilities[5]:

Stay in the origin country if:

$$(1-\rho)v \leq \left(\mu^{Mex}(X) - \mu^{US}(X) + k\right) + \frac{M + R - k}{\pi}$$

Migrate to the United States if:

$$(1-\rho)v > \left(\mu^{Mex}(X) - \mu^{US}(X) + k\right) + \frac{M + R - k}{\pi}$$

Return to origin country if:

$$\left(\mu^{Mex}(X) - \mu^{US}(X) + k\right) + \frac{M + R - k}{\pi} < (1-\rho)v \leq$$
$$\leq \left(\mu^{Mex}(X) - \mu^{US}(X) + k\right) - \frac{R}{1-\pi} - \varepsilon$$

The implications of the model might be seen easily with some graphs. I analyse the two reasons of coming back separately in order to see how the underlying motivation affects the selection of emigrants. Let's start by considering the case where temporary migration is only a matter of human capital gains and there is no uncertainty in the American market. In that case **ε** comes from a degenerated distribution with a mass point at ***ε=0***.

2.1 Model with human capital gains

The comparison of the return to unobserved ability in each place determines who migrates. Figure 1 represents the selection when unobserved ability is better paid in the United States than in Mexico (***ρ>1).*** In that case only high skill individuals migrate. Notice that the possibility of increasing someone's productivity by migrating temporarily incentives some people to

5. To make the investment in human capital interesting at all, the only assumption that they impose is *k>M+[R/(1-π)]*

migrate temporarily instead of staying at home. That is the reason why temporary migrants are placed in the middle of the other two groups.

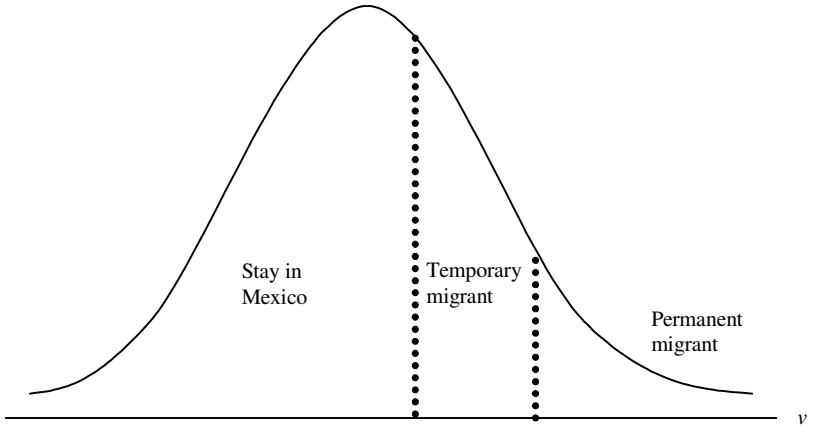


FIGURE 1

Figure 2 represents the opposite selection when unobserved ability is better paid in Mexico ($\rho<1$). In that case, only low skilled ability migrants migrate. Again, some people who otherwise would have stayed are willing to migrate with the option of doing it temporarily.

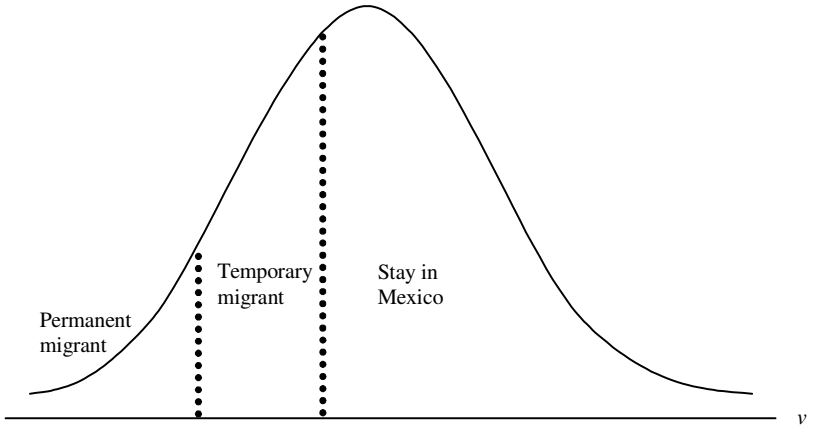


FIGURE 2

As Borjas and Bratsberg point out, returning migrants exacerbate the selection of migrants. This is important because in the presence of selection in terms of unobserved ability, the distribution of non-returning migrants' ability is closer to the returning migrants' distribution than to the distribution of non-migrants.

Moreover, the comparison of the mean ability component of returning migrants and non-migrants might be used to determine the type of selection of migration in unobserved terms since:

$E(v|nonmig) \leq E(v|retmig) \leq E(v|nonretmig)$

If the mean ability of returning migrants is higher than the mean ability of non-migrants there would be positive selection in unobserved human capital. Instead, if returning migrants have a lower ability than non-migrants there would be negative selection.

2.2 Model with human capital gains and uncertainty

With uncertainty the decision of migration does not change as long as ε is uncorrelated with v[6]. On the other hand, the set of migrants who decide to go back home changes. Indeed, although migrants leave with an idea of the time they want to stay abroad, once in the United States they might change changes plans: for instance there are people who planned to stay permanently that realise staying in the United States is not as good as they have thought, and there are people who planned to stay temporarily that finally decide to stay forever. As a result, the uncertainty moves the two distributions of returning and non-returning migrants closer, because now both supports coincide. This idea is better described with the following figures.

Figure 3 shows the decision of moving when returns to ability are higher in the United States than in Mexico. The first equation sets the cutoff above which high ability individuals migrate to the United States. When the actual wage in the United States is realised, people whose productivity is high enough decide to stay. It is clear from the figure how uncertainty works: since gains from migration are heterogeneous, people who are far away from the cutoff earn the most with the movement. Therefore, they are willing to stand very bad realisations of the productivity. Instead, it is fairly easy for a person close to the cutoff to draw a ε that makes him go back home. In a sense, those who are more willing to migrate are the ones who are less likely to go back.

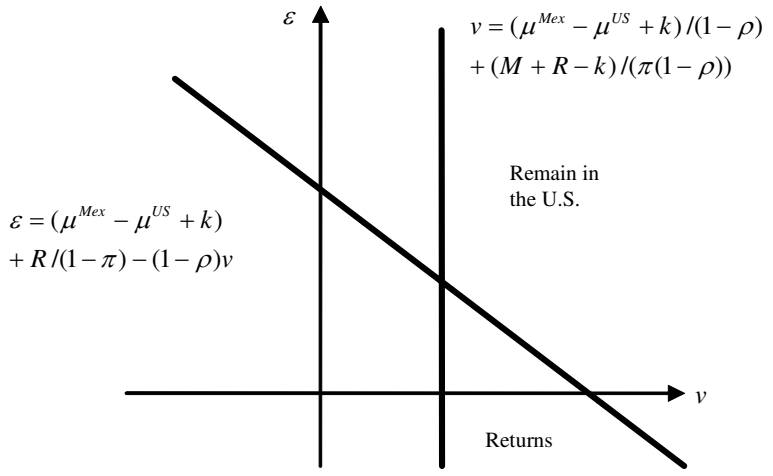


FIGURE 3

This result alone places returning migrants in the middle of the other two groups. This means that even in the case when migrants did not have human capital gains, those who were more willing to migrate would be less likely to go back home. Therefore, if there is

6. Even in the case of certain correlation many of the results will hold.

uncertainty, the usage of the distribution of abilities of returning migrants to estimate the distribution of non-returning migrants is also a good idea. Again the mean ability component of returning migrants might be used to specify the selection of migrants in terms of unobserved components.

Figure 4 shows the decision of migrating and returning when returns to ability are higher in Mexico than in the United States. In this case, only low skilled individuals migrate and again, those who are more willing to migrate will be less likely to go back.

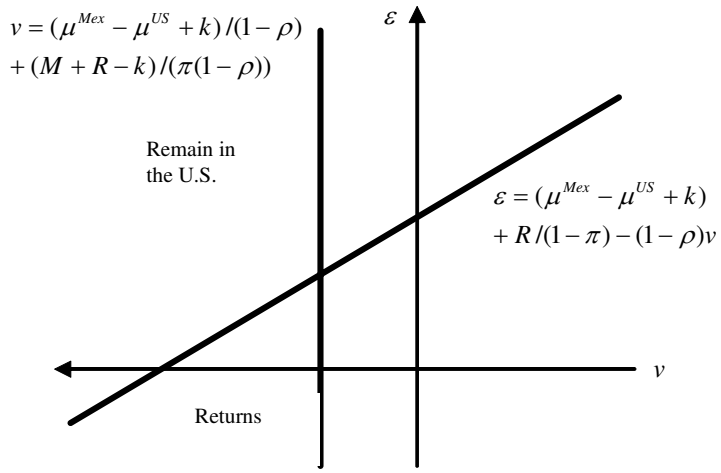


FIGURE 4

Concluding, Borjas and Bratsberg show that when there is selection in terms of unobserved productivity, the distribution of abilities of permanent migrants is much more similar to the distribution of abilities of temporary migrants than to the distribution of non-migrants. Since it is impossible to obtain testable implications for a model based in unobserved components of the productivity, the next section extends it with observed characteristics.

3 Testable implications

One idea that comes out from the model is that those individuals who are more willing to migrate in the ability dimension are also less likely to return. This implication cannot be tested for unobserved characteristics, but we may extrapolate the same reasoning to other observed dimensions. The higher the correlation of those characteristics with unobserved productivity, the higher should be the likelihood that the two probabilities go in opposite directions. That is the reason why I put special attention on two important observed components of human capital that should be correlated somehow with unobserved heterogeneity: education and age. However, I would like to check whether other characteristics that are less correlated with v behave the same way. In the appendix A I extend the model adding one observed characteristic independent respect to unobserved ability. In that model I check whether the probability of migration and the probability of returning still go in opposite directions. In general, those probabilities do not vary in opposite directions. However, some additional mild restrictions on the distribution of the unobserved components achieve this result. That is the reason why I present the test for many other dimensions.

3.1 Data

I use microdata from a 10% sample of the Mexican census to describe non-migrants[7]. I include as a non-migrant a person who did not migrate to the United States within the last five years.

I use microdata from a 5% sample of the American census to describe non-returning migrants. The American census does not ask anything related to legal status, therefore it should capture people who is immigrant alien (those with a valid visa of residence), temporary visitors (students, temporary workers or exchange program individuals) and undocumented. Regarding temporary visitors, the census is not expected to capture tourists, very short term visitors and commuters because individuals who live or stay at another place most of the time should not fill out the questionnaire. On the other hand, some illegal immigrants may have a high incentive to hide their situation in the US Lacuesta (2005) examines more carefully this issue using residual techniques as in Heer and Passel (1987) and death reports as in Borjas et al. (1991). This analysis determines that underreporting is not overwhelmingly biased according to characteristics.

I use the Mexican census to describe returning migrants. There are three types of returning migrants that have plans to work abroad: commuters, seasonal workers and temporary workers. In this paper I do not treat the first two groups. These two groups do not fit well in the model that was sketched in the previous section. Commuters go back and forth from the US to Mexico. They usually live in the border states of Mexico and go to the United States to work on a daily or weekly base. Because of the question in the Mexican census, they should not appear as returning migrants. Seasonal workers are Mexicans who go to the United States every year. They work in Mexico part of the year, and they work in the United States during a peak season in particular sectors. Agriculture is the typical example for this type of worker. This sector has important seasonal peaks followed by periods of unemployment that make the population willing to accept other jobs elsewhere. The Mexican census is likely to pick those individuals as returning migrants. In order to drop this group,

7. For a brief description of the data sources go to the Appendix B.

I consider the pool of migrants who stayed in the US longer than three months. In principle, seasonal jobs are expected to be held at most for three months. In order to check the robustness of the analysis, I consider two alternative samples that should also drop seasonal workers. The first one is restricting the sample to migrants who migrate between 1995 and 1996. A seasonal worker migrates on a yearly basis; therefore, seasonal workers migrating in 1995-1996 will probably leave again the following years. Since in the Mexican census people only report their last trip, those individuals will not show up in that pool. The second one is to restrict the set of migrants to individuals who departed from states outside the border. An individual from the border is likely to have more information about the seasonal peaks in the United States, therefore, it is expected that most of the seasonal workers migrate from bordering regions. All samples produce qualitatively similar results but some variations in terms of magnitudes, when this is the case all samples are reported.

Table 1 quantifies the stock of migrants in 2000 and the flow between 1995 and 2000 using the two different sources. All numbers are expanded using the factors that each agency provides.

Table 1: Some Magnitudes about Mexican Migration in the U.S. and Sources

		Mexican Population	Mexican Immigrants in the U.S.	Flow 95-99 of Mexican non-returning	Flow 95-99 of Mexican returning by 2000	Flow 95-9 of Mexican returning by 2000 staying more than 3 months
Males						
	Mexican Census	46.640.360			200.044	185.572
	American Census		5.166.943	1.396.701		
Females						
	Mexican Census	49.152.676			56.251	47.923
	American Census		4.158.509	1.023.399		

Source: Mexican and American Census 2000

The following analysis is restricted to males. There are two main reasons for imposing this restriction: on the one hand, male immigration is quantitatively more important. Therefore, when using the sample of returning migrants, I have a big number of observations that increases the confidence in the results. On the other hand, the model explicitly treated economic factors underlying migration and returning migration. However, a big fraction of the female movements are still motivated by the reunification of the household [Bustamante (1997)].

3.2 Empirical results

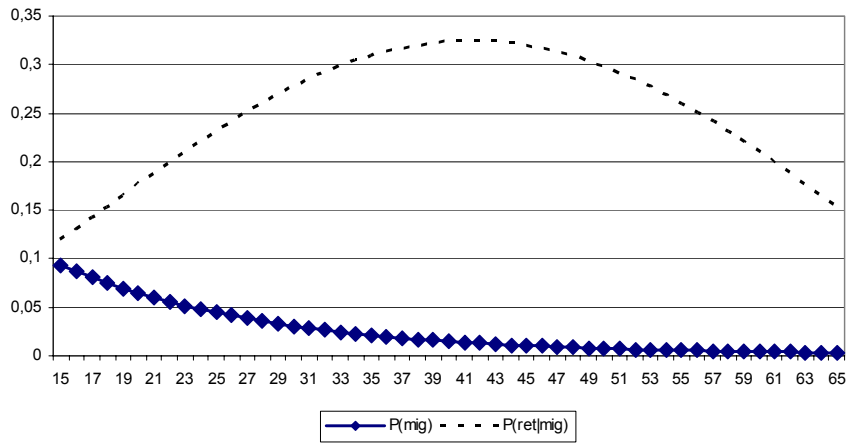
According to the previous model, the selection of migrants and returning migrants should go in opposite directions. In this section we provide some evidence about this prediction.

The table in the Appendix C presents the results of two different probits: one for the decision of migrating and another for the decision of coming back[8]. Panel A shows the analysis using individual characteristics. The first column of Panel A shows the migration decision between 1995 and 2000. All male migrants within this period are used regardless their returning status. On the other hand, the second column analyses the returning decision for those migrants who stayed abroad more than 3 months.

8. Results appear to be robust to the usage of other non-parametric estimations of the probability. Actually, straight tabulations of raw data capture the same shape described with the coefficients of the probit.

The two probits include regressors that proxy human and social capital in 1995[9]. Panel A shows that most of the variables affect differently the probability of emigration and the probability of coming back. Let's focus on the two variables that better describe the human capital of a person: age and education. Figures 5 and 6 show how age and education affect the two probabilities respectively.

Figure 5: Estimated probability by age



Age enters negatively in the probability of migration. If assimilation in the American labour market is higher for young migrants as DeLeire et al. (2005) point out, returns to emigration will be higher at young ages, and older individuals would be less willing to migrate. This result has been overwhelmingly found in the literature [Massey and Espinosa (1997), Hanson and Chiquiar (2005)].

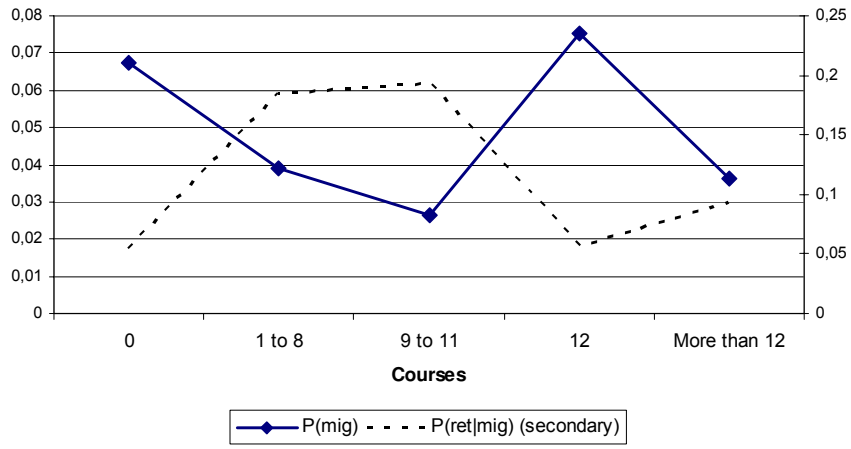
On the other hand, the probability of returning given that someone migrated is increasing with age until the age of 43 and it decreases from that age onward[10]. This downward pattern at high ages might be attributed to an increasing proportion of tie movers.

As already shown by Hanson and Chiquiar (2005), the rate of migration is non-monotonic across educational levels[11] (see figure 6). Migrants typically depart from rural areas located in the middle of Mexico. Residents in those areas are generally less educated; therefore, failure to control for the region of birth may generate part of the non-monotonicity by increasing the coefficient of very low educated individuals (0 courses). However, if each observation in the Mexican census is weighted in such a way that regions and sizes of villages reproduce the regions of departure of migrants, the same non-monotonicity is kept[12].

9. Age and its square as well as some dummies for educational level are used as an approximation for human capital. Marriage and family ties at destination are considered as social capital in Mexico and in the United States respectively.
10. This concave pattern was already found in DaVanzo (1976) and Massey and Espinosa (1997).
11. This educational sorting is pretty constant across time. When the same computations are done using data for 1960 and 1990 the same sorting is obtained.
12. The region of origin is not specified in the American census. By using information about regions of departure of Mexicans who migrated and did not come back from the Mexican census we may re-weight observations of non-migrants by: [(P[region|mig]) / (P([region|non-mig])].

There are few papers analysing the educational sorting of returning migrants and most of them restrict the effect of education on the probability to be linear[13]. I use several dummies for the educational levels instead, and the specification turns out to be relevant. Actually, there is also a non-monotonicity in the probability of returning by educational attainment, and very interestingly, this non-monotonicity is symmetric to the one obtained in the probability of migration.

Figure 6: Estimated probability by education



The probit specification includes some other controls to proxy social capital in both countries. Marriage captures a higher attachment to Mexico. The sign of this variable in the probit is positive. However, the interpretation of marriage as a higher attachment to Mexico would suggest a different sign. In the future, we may want to incorporate other household considerations in order to explain this behaviour, for example marriage may increase the household's economic needs. In that case, it would not be surprising that the coefficient of the probit was positive. The coefficient on marriage is also reversed when analysing the probability of coming back given that the individual migrated. However, the sign is not significant.

Previous family migration captures social capital in the destination country. It is constructed by looking at other migrants within the household who migrated in 1995. For those individuals in the US the dummy variable takes the value 1 if there is any other member in the household who migrated during that year[14]. The sign of the coefficient is positive as expected. The coefficient of family ties at destination is also reversed when considering the probability of returning.

Since individual data only provides four dimensions of analysis, I analyse the behaviour of the coefficients on the two probabilities when household characteristics are incorporated. Panel B in Appendix B shows the analysis using household characteristics.

---

13. In particular different results were found. Some researchers argue that the odds of coming back increases with education [Long and Hansen (1975) and Lindstrom (1996)], whereas other indicate the opposite [Massey and Espinosa (1997), Beenstock (1996)].

14. I used other specifications for this variable. One alternative is "any person known in the American market". It would take 1 in the Mexican census if the household receives some remittances from abroad. It would take 1 in the US if the Mexican lives with other Mexicans who arrived before.

The data for this probit comes from the Mexican census alone. I can only include household characteristics that are affected in a known way by the migration[15]. Mindful of this problem, I include as regressors: the size of the household[16], the number of children[17], the number of elderly[18], the number of members working and the number of members studying. Finally, I also include some information about resources of the household apart from wages and international remittances and past migrations in the household. Notice that since the information comes from the Mexican census, I can control for the region of residence and the size of the village.

It is beyond the scope of this paper to interpret these coefficients; however, a quick look at the signs in column (1) and column (2) in Panel B shows that 6 out of 7 coefficients go in opposite directions. This evidence reinforces the idea, as it was derived from the model in section 2, that returning migrants are in terms of characteristics between non-migrants and non-returning migrants.

15. I only know gender and age of the migrant but I do not know for example whether the migrant was the head of the household or his educational level.

16. The migrant is added as a member of the household.

17. A child is an individual below 15.

18. An elderly member is above 65.

4 Effects of emigration on returning migrants

The literature on benefits of emigration in the country of origin has overwhelming treated remittances as the main gain for those left behind. However there could be an additional source of welfare gains derived from the new skills that returning migrants may have learnt from their experience abroad. Following this idea, I study the possibility that emigration abroad increases someone's productivity upon return. There are two channels by which human capital gains derived from emigration may affect someone's labour behaviour in Mexico. On the one hand, the individual might upgrade his occupational status. On the other hand, the individual might increase his wage at his previous job upon return. The second channel is what economists usually prioritise, but there are several reasons to study these two steps one by one. First, we only have wages of those who previously decided to be wage-earners, instead of being out of the labour force or being self-employed. We might assume that this decision is more or less similar for migrants and non-migrants. However, the selection problem would be exacerbated if emigration affects the decision of participating. Actually, there are reasons to believe that this might be the case because migrants might increase savings that are useful to either retire or to set up a business, therefore this test seems a natural first step before running any wage regression. Additionally, occupational changes might be a proxy of wage increases since there are sectors that are better paid than others. The nice thing about occupational changes instead of wage increases is that I benefit from the availability of a unique data set with the whole occupational career of migrants and non-migrants from rural areas in Mexico (Mexican Migration Project that will be described in the following sub-section).

4.1 Does migration change someone's labour status?

Table 2 provides information about the sectors where returning migrants and non-migrants worked in the year 2000. The sample is restricted to males between 25 and 50 years old. Returning migrants are more likely to be employed in agriculture and construction whereas non-migrants are more likely to be employed in services. In mean, the sectors where migrants work used to pay lower wages; therefore their status after migration appears to be relatively bad respect to the non-migrant population. However with the information in this table we cannot say anything about their situation before migration and any possible upgrading derived from emigration.

On the other hand, the percentage of males between the age of 25 and 50 who report being out of the labour force is smaller than 3% for returning migrants and non-migrants. Both percentages are similar and not statistically different. The difference between the two groups increases over time. In particular, when we look at individuals between 55 and 65, returning migrants are slightly more likely to be out of the labour force. This could mean that some of them decide to retire right after their return. The difference is however not economically significant.

Finally, it can also be seen in Table 2 that, after migration, returning migrants are more likely to be self-employed than non-migrants. The difference in the probability of being self-employed between the two groups is about 8%. This differential can also be seen for other cohorts and it increases with age. As it happened before with the differences in occupations, we do not know whether the business of the migrant was already set up before the movement or the individual set it up afterwards.

Table 2: Labor status of non-migrants and returning migrants. Males between 25 and 50

	Non-migrants	Return Migrants
Out of the labor force (%)	1,99%	2,90%
Self-Employed (%)*	29,87%	37,73%
Sector		
Agr., Forest., fish., hun.	18,35%	27,62%
Mining	1,19%	1,12%
Utilities	0,93%	0,46%
Construction	13,22%	22,90%
Manufacturing	20,72%	18,27%
Wholesale trade	1,41%	1,16%
Retail trade	8,65%	6,38%
Transp. and warehousing	6,99%	5,62%
Information	1,11%	0,75%
Finance and insurance	1,02%	0,53%
Real state, rental, leasing	0,28%	0,22%
Prof., scientific and tech. services	1,52%	0,92%
Management of cies & enterp.	0,03%	0,02%
Administrative and support	2,08%	1,01%
Educational services	6,10%	1,97%
Health care and social assistance	2,07%	0,77%
Arts, entertainment, and recreation	0,65%	0,53%
Accomodations and food services	2,51%	1,73%
Other services (no public adm.)	4,97%	3,42%
Public administration	6,21%	4,61%

Source: Mexican Census 2000
* Self-employed include those who report to be self-employed, working in a business of the family or being "patron"

In order to shed some more light about potential occupational changes before and after migration I use another data set. I use information from the Mexican Migration Project [MMP described in Massey and Zenteno (2000)]. In this data set, some heads of the household of particular rural areas in Mexico were asked about their migration experience and their labour careers[19]. The MMP has no information about prior wages, only wages after migration and they are difficult to match with a particular year. Instead, the survey has very valuable information about occupations held by an individual during his whole life.

I track the occupations for a specific cohort of individuals (workers born between 1950 and 1955). I split the sample in two groups (returning migrants and non-migrants), and I consider year by year the percentage of people in each group that are working in a particular sector in Mexico[20].

People migrate around the age of 20 and come back home around the age of 27. Any difference between the occupations of returning migrants and non-migrants before the ages of 16 and 22 reflect pre-migration differences that cannot be attributed to the departure. Instead, if emigration has an effect in the occupational status, those previous differences should change after migration as a consequence of the move. Figures 7 to 11 graph the evolution of occupations for both groups.

19. One shortcoming of this data set is that I cannot extrapolate the conclusions to better educated individuals because most of the communities are in rural areas without a sizeable number of highly educated individuals.
20. The group of temporary migrants in a particular year is formed by those who are residing in Mexico in that particular year.

Migrants are more likely to be working at young ages. As it was seen in Table 2, the proportion of migrants working in agriculture is large (between 30% and 40%) and higher than the proportion of non-migrants. Migrants also have a higher percentage of workers in manufacturing and self-employment, although the difference respect to the non-migrant population is very small for every age[21]. On the other hand, many more non-migrants are working in services compare to migrants[22].

Figure 7: Percentage out of the labor force during the life cycle

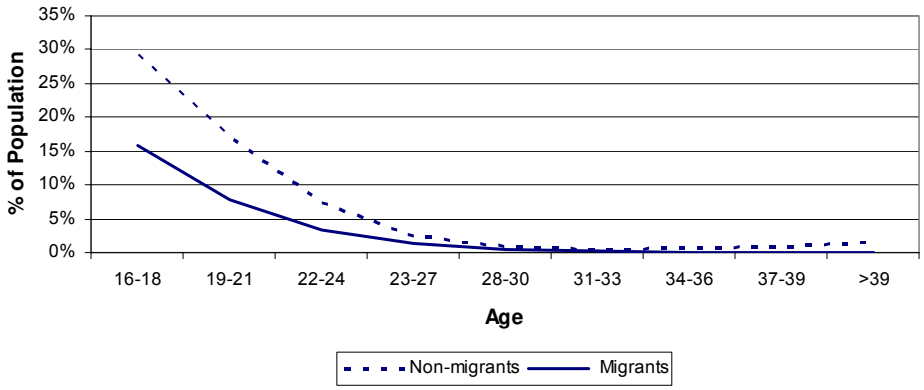
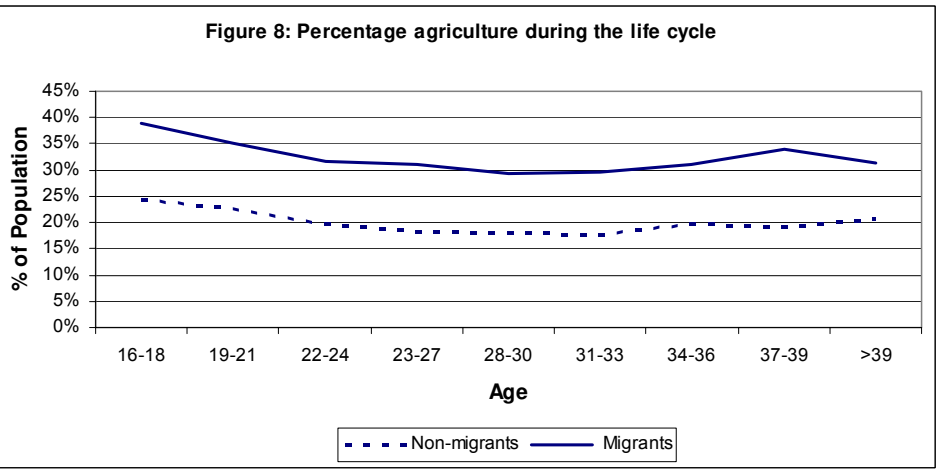
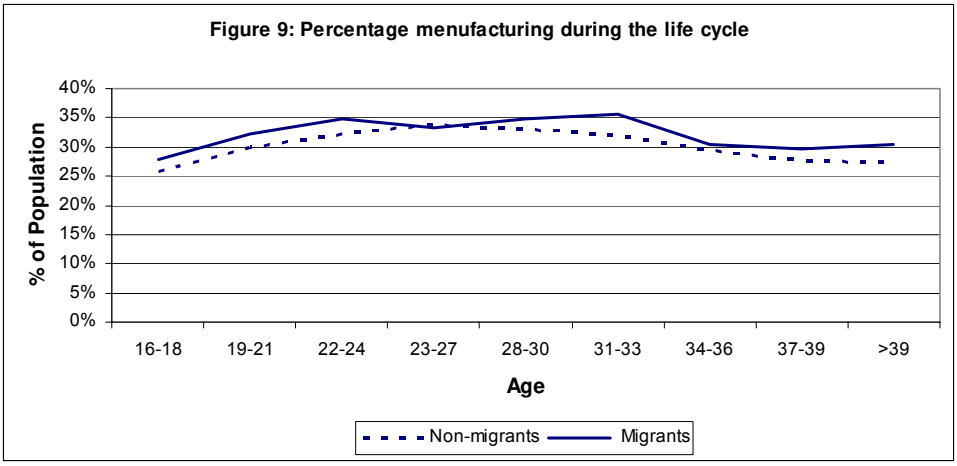


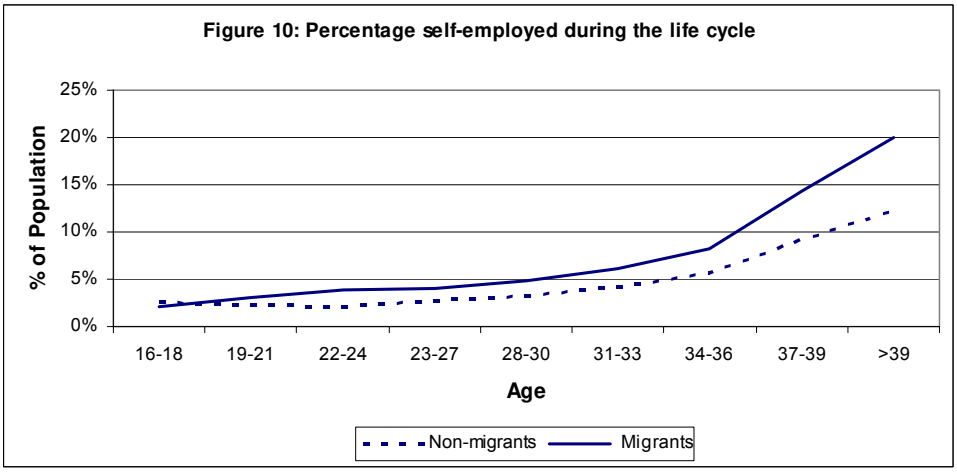
Figure 8: Percentage agriculture during the life cycle

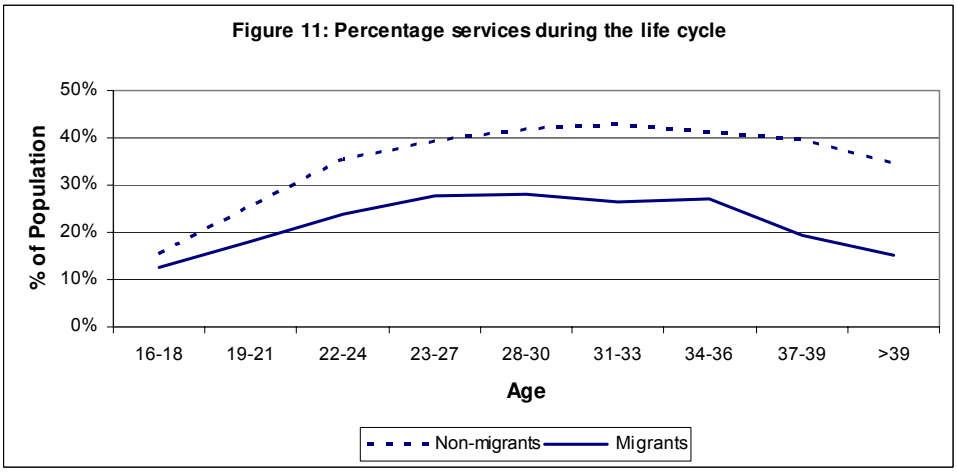


21. For some cohorts, migrants are more likely to be self-employed whereas for others, the difference between the two groups is almost imperceptible.

22. This is also robust to different cohorts.







In terms of changes over the life cycle, non-migrants usually enter into the labour force at later ages. However, after the age of 25, both proportions remain similar and it does not appear that migrants are leaving the labour force at a higher likelihood than non-migrants. This is consistent for different cohorts.

In terms of agriculture and manufacturing, the difference between migrants and non-migrants' proportions is kept constant for the whole life profile. This pattern is robust for different cohorts. This would suggest that migration has no effect in the probability of leaving/entering agriculture or manufacturing.

In terms of self-employment the difference observed during the life cycle between migrants and non-migrants increases slightly for the last two year groups (over age 37). The difference between those two groups before the age of 36 was about 2%. Afterwards, this difference increases to 8%. This effect, as it could be seen in figure 7 comes from a greater decrease of migrants in services. Migration spells usually occur before the age of 37; therefore, the evidence would suggest that there could be an effect of migration on the probability of setting up a business. However, since migration mainly occurs in the early 20's, this change does not occur just after the return.

This preliminary evidence suggests that migrants come from a certain type of sectors. They overwhelming stay at the same sector after migration although there is some evidence that migration may affect self-employment. Nevertheless, the effect does not appear to occur just after the migration.

4.2 Human capital gains within particular sectors

Borjas and Bratsberg (1996) identify temporary migration as a mean to increase labour productivity. In the previous sub-section I showed some evidence that emigrants upon return usually work in the same sector they were working before migrating, at least in the short run. However, the previous evidence was not useful to analyse human capital gains that occurred during the emigration and that increase the productivity in the sector the migrant was working before migrating. In order to check whether there is a premium at all, I run Mincer regressions pooling non-migrant and returning migrants. The regression includes human capital variables (age, age square and education) as well as region and dummies for the size of the village in order to eliminate differential wages across locations in Mexico. I also include sector dummies in order to see wage differences within a particular occupation. The coefficient of the dummy "retuning migration" identifies the wage gap of this particular group of workers with respect to non-migrants that cannot be explained by observed characteristics.

Table 3 shows that returning migrants earn significantly more than observationally equivalent non-migrants. For all the specifications of returning migrants the premium is positive. It is above 7% when returning migrants is defined as the group of migrants staying more than 3 months and it is around 10% for the other specifications.

Table 3: Wage regression pooling return migrants and Mexican residents

	>3 months	Migrants 95-96	Outside the border
Migration	0,0755	0,1234	0,0974
	(0,0027)**	(0,0055)**	(0,0027)**
Age	0,0371	0,0371	0,0363
	(0,0001)**	(0,0001)**	(0,0001)**
Age Square	-0,0004	-0,0004	-0,0004
	(0,0000)**	(0,0000)**	(0,0000)**
1-8 courses	0,1456	0,1457	0,142
	(0,0008)**	(0,0008)**	(0,0008)**
9-11 courses	0,303	0,3034	0,3
	(0,0008)**	(0,0008)**	(0,0009)**
12 courses	0,5309	0,5313	0,5323
	(0,0010)**	(0,0010)**	(0,0011)**
University or more	1,1523	1,1521	1,1527
	(0,0011)**	(0,0011)**	(0,0012)**
Region dummies	Yes	Yes	Yes
Size of the Village	Yes	Yes	Yes
Sector	Yes	Yes	Yes
Constant	4,0159	4,0167	4,0092
	(0,0024)**	(0,0024)**	(0,0026)**
Unweighted observations	1.102.067	1.102.067	903.804
Return Migrants	5.456	1.230	5.302
R-squared	0,44	0,44	0,45

Robust standard errors in parentheses
* significant at 5%; ** significant at 1%

Source: Mexican Census 2000 weighted by the factor provided by INEGI
Sample restricted to males between 15 and 65 in 1995 not self-employed who have positive experience and who are not studying
Omitted dummy 0 courses

There are two interpretations for the wage differential coming directly from the equations in section 2. The easiest one is that emigration increases the productivity of individuals upon return. On the other hand, this wage gap may refer to different ability components of migrants and non-migrants that already existed before migration. With a single cross section is impossible to identify these two components unless we add some additional restrictions. In order to highlight this point notice that the model in section 2 specifies the wage in Mexico as a function of characteristics *X* and ability *v*. Additionally if someone migrates and decides to come back, he will earn a premium of k. With this model the wage equation will be:

$$\ln w_i = \mu_0 + \mu^{Mex}(X_i) + k1(ret\_mig_i) + v_i$$

We are going to estimate with OLS the following regression:

$$\ln w_i = \alpha_0 + \beta X_i + \gamma 1(ret\_mig_i) + \omega_i$$

If we take expectations of the real model we shed some light on the identification of the parameters:

$$E(\ln w \mid X, 1(ret\_mig)) = \left[\mu_0 + E(v \mid non\_mig, X_i)\right] + \mu^{Mex}(X_i) + \\ + 1(ret\_mig_i)\left[k + E(\varepsilon_i \mid ret\_mig_i, X_i) - E(\varepsilon_i \mid non\_mig_i, X_i)\right]$$

Actually the correlation between the error term and the dummy for migration will generate that:

$$p\lim \hat{\gamma}_{OLS} = k + E(v \mid ret\_mig, X) - E(v \mid non\_mig, X)$$

If *E(v)=0,* and assuming that the *P(retmig|X)=p(X)*

$$p\lim \hat{\gamma}_{OLS} = k + \frac{E(\varepsilon \mid ret\_mig, X)}{1 - p(X)}$$

In order to identify the two components with a single cross section and without the availability of good instruments, I must impose some additional restrictions about the way the two components vary among individuals. Actually the model in section 2 presents a very neat way of identifying these two components. The model exposes two reasons to go back: human capital gains and non-realised expectations of someone's productivity abroad. Human capital gains require some time to accumulate. If the realisation of the productivity in the United States is too poor, there is no incentive to stay in the United States and it is better to go back to Mexico immediately even without the accumulation of the human capital gains. In terms of the equations of section 2, someone will stay temporarily if:

$$\varepsilon < \mu^{Mex}(X) - \mu^{US}(X) + k - \frac{R}{1-\pi} - (1-\rho)v$$

Moreover, the person will go back immediately if:

$$\varepsilon < \mu^{Mex}(X) - \mu^{US}(X) - \frac{1-\pi}{\pi}k - (1-\rho)v$$

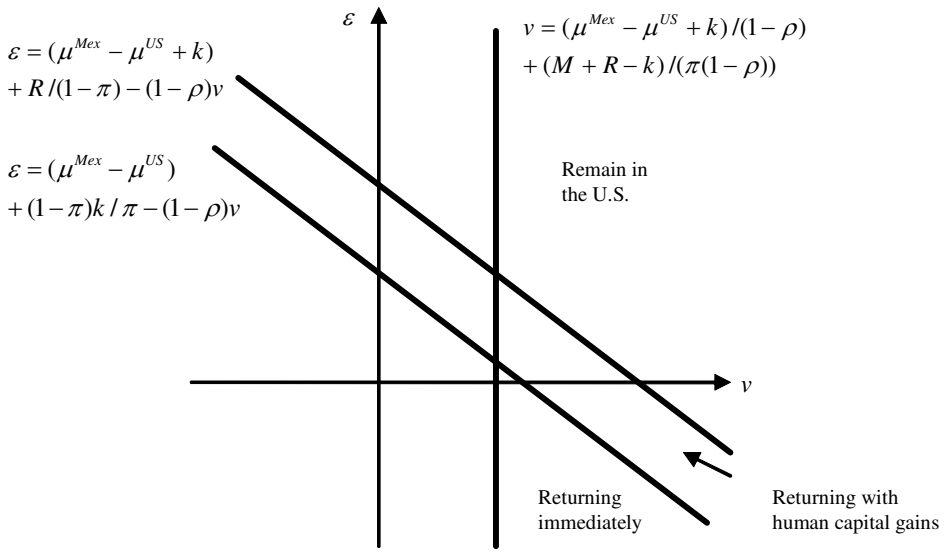


FIGURE 12

The conditions stated before are enough to ensure that the three possible situations will arise together: some people will come back immediately, some people will stay temporarily and some people will not come back. Figure 12 represents the new sorting. There is a new line respect to those in figures 3 and 4 that represents the set of individuals who come back home immediately since their productivity in the United States was really low. The figure is instructive to point out that those individuals are again selected in terms of ability; actually it can be shown that:

$E(v|non\_mig) \leq E(v|ret\_mig\_immediate) \leq E(v|ret\_mig\_temp) \leq E(v|non\_ret\_mig)$

Notice that the model departs from the hypothesis that individuals discover immediately their productivity in the United States. This is not necessary as long as the process of discovering someone's productivity is independent of the unobserved ability.

As an example of how to identify the two components with a simple regression let's say that someone requires one year of migration to increase the human capital, then a regression such as the following would identify the importance of the two components:

$$\ln w_i = \alpha_0 + \beta X_i + \gamma_1 1(ret\_mig_i) + \gamma_2 1(ret\_mig_i) 1(mig > 1 year_i) + \omega_i$$

Let's take expectations of the true model:

$$\begin{aligned} E(\ln w \mid X, 1(ret\_mig), 1(mig > 1\_year)) = &\left[\mu_0 + E(v \mid non\_mig, X)\right] + \mu^{Mex}(X_i) + \\ &+ 1(ret\_mig_i)\left[E(v \mid ret\_mig\_immed, X) - E(v \mid non\_mig, X)\right] + \\ &+ 1(ret\_mig_i) 1(mig_i > 1\_year)\left[k + E(v \mid ret\_mig\_temp, X) - E(v \mid ret\_mig\_immed, X)\right] \end{aligned}$$

$$\begin{aligned} p\lim \hat{\gamma}_{1,OLS} &= E(v \mid ret\_mig\_immed, X) - E(v \mid non\_mig, X) \\ p\lim \hat{\gamma}_{2,OLS} &= k + E(v \mid ret\_mig\_temp, X) - E(v \mid ret\_mig\_immed, X) > \gamma_1 \end{aligned}$$

With this type of regression we can evaluate the importance of the unobserved selection component respect to the human capital gain for a particular group of migrants. Table 4 reports those wage gaps for the three specifications of returning migrants[23]. In all cases the wage gap is large for the shortest duration in the United States and it does not increase statistically with the length of the trip. The results are kept if the shortest length of the trip is less than 6 months. The results appear to be consistent with a story without human capital gains where the information about someone's productivity is realised fairly soon. However we must interpret this evidence with caution since it is impossible to know whether there is an immediate reward to American experience. If this was the case, even for the shortest spells we would not be able to identify the two components.

23. The smallest number of returning observations for a particular length is 169.

Table 4: Log Wage Differential of Returning migrants and non-migrants by duration of the migration

	>3 months	Migrants 95-96	Outside the border
Migration	0,0787	0,1096	0,1126
	(0,0135)**	(0,0420)**	(0,0133)**
Duration between 1 and 2 years	-0,025	0,0432	-0,0573
	(0,023)	(0,061)	(0.0239)*
Duration between 2 and 3 years	0,0184	0,0458	-0,0287
	(0,0368)	(0,0637)	(0,0367)
Duration more than 3 years	0,0176	-0,0111	-0,0036
	(0,0403)	(0,0567)	(0,0421)

Source: Mexican Census 2000. Same regressions as Table 3 with interactions of migration and durations
Sample restricted to males between 15 and 65 in 1995 not self-employed who have positive experience and who are not studying
Omitted dummy 0 grades

Summarising this evidence reinforces the idea that migrants are positively selected in terms of unobserved ability. Actually the lower bound for the mean ability component of those migrants is very high. This increases the importance of taking into account the distribution of those unobserved components when estimating the productivity of those individuals who are abroad. Moreover, human capital gains derived from emigration does not appear to be very important. This evidence will be very useful in order to estimate the productivity of those migrants who are still abroad, because wages won't be contaminated with human capital gains.

5 Human capital loss in Mexico derived from emigration

As it was said in the introduction, Hanson and Chiquiar (2005) estimated the counterfactual wage of Mexican immigrants in the United States using the following formula:

$$g_{mig}(w^{Mex}) = \int \theta f_{non\_mig}(w^{Mex} \mid x) h_{non\_mig}(x) dx$$

where θ is the re-weighting function that expresses the different mix of characteristics for each particular group.

$$\theta = \frac{\Pr(i = mig \mid x)}{\Pr(i = non\_mig \mid x)}$$

The model in section 2 showed that the main problem of this estimation is the self-selection of migrants in terms of unobserved ability that produces a difference between the wage distribution of migrants and non-migrants $f_{mig}(w^{Mex} \mid x) \neq f_{non\_mig}(w^{Mex} \mid x)$. Moreover, the model proposed using the information of returning migrants to improve the estimation of non-returning migrants' ability. Therefore, I am going to use returning migrants' wages $f_{ret\_mig}(w^{Mex} \mid x)$ and the following weighting factor:

$$\frac{\Pr(i = non\_ret \mid x, mig)}{\Pr(i = ret\_mig \mid x, mig)}$$

I incorporate two additional changes respect to Hanson and Chiquiar's estimation:

1) I take into account the different region of residence in Mexico for returning and non-returning migrants. For simplicity I assume that all the other observable characteristics are exogenous to the residence once controlling by migration status:

$$\frac{\Pr(i = ent \mid x, non\_ret)}{\Pr(i = ent \mid x, ret\_mig)} = \frac{\Pr(i = ent \mid non\_ret)}{\Pr(i = ent \mid ret\_mig)}$$

2) In order to be left with a better approximation of the human capital that is lost by Mexico, I drop the component of salaries that is due to purchasing power differences at the regional level.

$$\ln w_i^{Mex} = \alpha_0 + \beta X_i + d_r region_i + d_r size\_village_i + u_i$$
$$\hat{u}_i^{Mex} = \ln w_i^{Mex} - \hat{d}_r region_i - \hat{d}_r size\_village_i$$

Hence, the final estimation of the human capital that is lost is:

$$g_{non\_ret}(\hat{u}_i^{Mex}) = \int \theta^{final} f_{ret\_mig}(\hat{u}_i^{Mex} \mid x) h_{ret\_mig}(x) dx$$

$$\theta^{final} = \frac{\Pr(i = ent \mid non\_ret)}{\Pr(i = ent \mid ret\_mig)} \frac{\Pr(i = non\_ret \mid x, mig)}{\Pr(i = ret\_mig \mid x, mig)}$$

The kernel used for the estimation is Gaussian with a bandwidth of 0,07 as in Hanson and Chiquiar (2005)[24]. Figure 13 replicates the methodology of Hanson and Chiquiar using data in 2000. They used the wage of non-migrants to estimate the potential wage in Mexico of migrants. I obtain similar results to what they found with the 1990 data: the distribution of human capital of migrants and non-migrants are very similar. However, migrants are slightly more likely to be drawn from the middle of the distribution. In figure 14 I take into consideration the region of departure. It is observed that this factor does not produce many differences, although respect to Hanson and Chiquiar's distribution, there is a tendency that makes migrants still more likely to be drawn from the middle part of the distribution. This is due to the fact that migrants come from regions with an intermediate level of income.

Finally, figure 15 shows the change in the estimation when the information of returning migrants is used. Respect to Hanson and Chiquiar's estimation, the usage of returning migrants' wages instead of non-migrants' wages clearly draws some people from the bottom part of the distribution and adds them to the middle and upper part of it. This is so because of the positive premium of returning migrants with respect to non-migrants. The upper part of the distribution does not change much because there are not many old and highly educated migrants; however, the very top part of the estimation increases slightly. Since returning migrants are more concentrated than non-migrants, the estimation of abilities for non-returning migrants is also more concentrated. A new estimation of the mean human capital of migrants would not change much respect to what the model without unobserved heterogeneity would predict, however, the variance and the different percentiles would change a lot. Therefore, failing to control for unobserved factors may contribute to enormous biases to answer relevant policy questions in terms of distributions.

24. In the estimation of wages, the constant refers to the wage in a village of Michoacan of less than 2.500 inhabitants. This is the location where migration was higher.

Figure 13: Distribution of log-wages of migrants

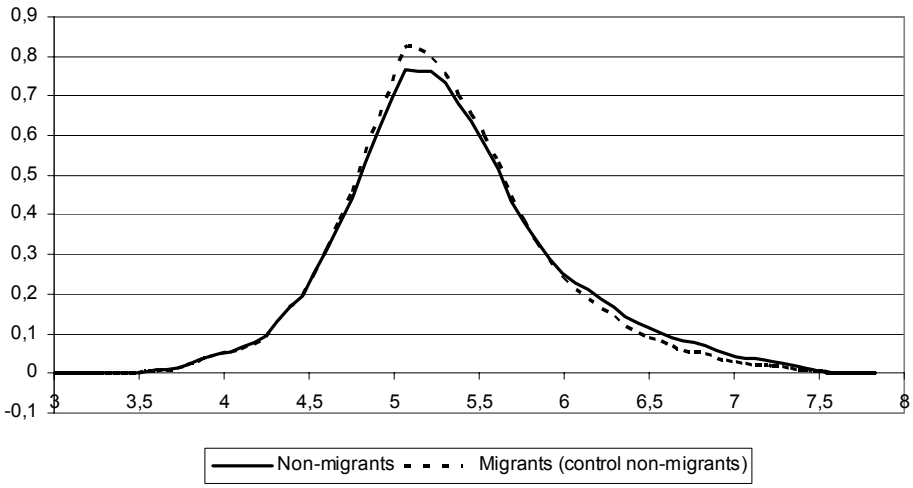


Figure 14: Distribution of log-wages of migrants

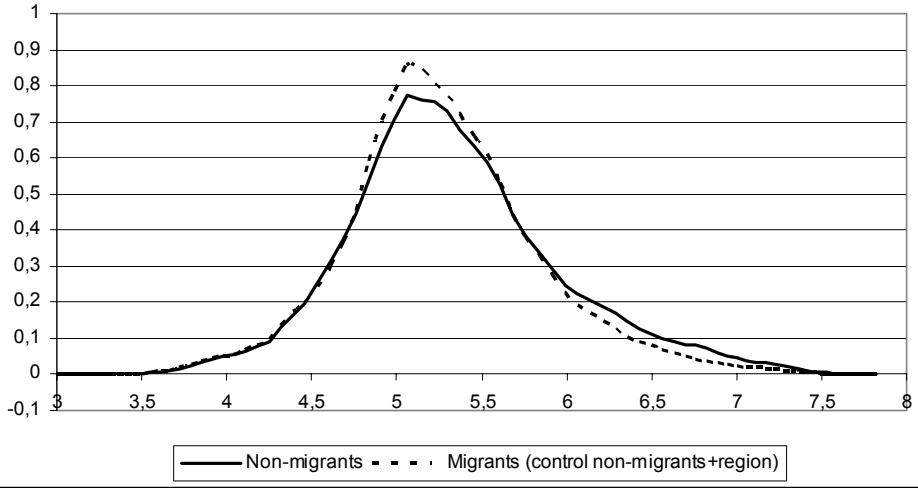
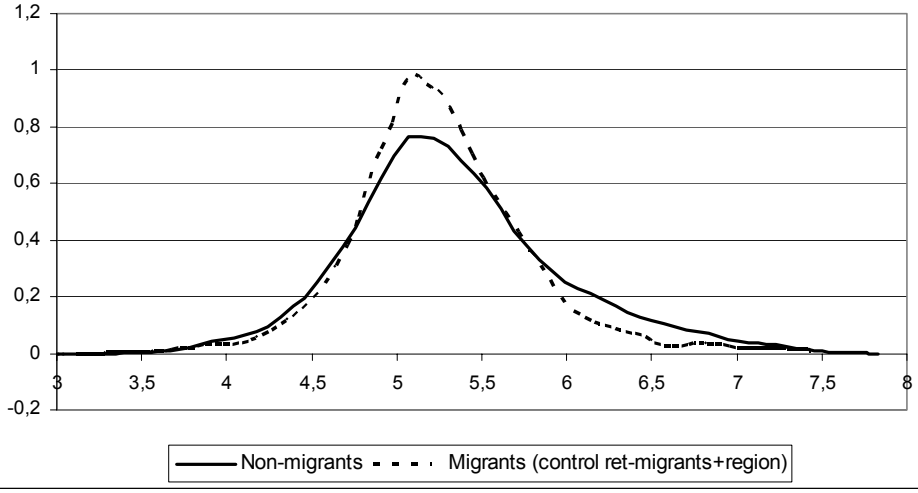


Figure 15: Distribution of log-wages of migrants



6 Conclusions

In this paper I have studied international migration with the ultimate goal of estimating the human capital loss that emigration generates in Mexico. Unlike other studies, I consider the selection of migrants in terms of observed characteristics as well as unobserved factors that affect productivity at home. Since census data do not provide information about the wage in Mexico of migrants before their departure, we need to look for an alternative group of residents in Mexico similar to permanent migrants in terms of productivity. The paper analyses the use of returning migrants' wages for estimating the productivity in Mexico of non-returning migrants. The analysis follows three steps: presentation of a model with the selection of migrants, non-migrants and temporary migrants; verification of some important implications of the model in the data; and analysis of potential human capital gains derived from emigration. First, I use the model by Borjas and Bratsberg (1986) to analyse the selection of permanent migrants, returning migrants and non-migrants. This model claims that returning migrants lie between non-returning and non-migrants in terms of unobserved abilities. Moreover, the paper shows that a crucial implication of the model holds in the data, namely that the probability of migration and the probability of coming back go in opposite directions in several dimensions. However, in order to use the wage of temporary migrants as a proper estimator of the productivity of non-returning migrants, we need to reject any human capital gain generated during the migration. The paper shows two pieces of evidence against the idea of human capital gains derived from emigration. On one hand, using information about the sector where Mexicans worked during their life cycle, the paper shows that migrants do not overwhelmingly change their labour status right after their migration. On the other, although temporary migrants earn significantly more than non-migrants, this difference is already evident even when comparing non-migrants with migrants who stay a very short period of time in the United States. This would mean that even those individuals who did not have time to acquire any human capital gain present a wage differential in relation to non-migrants.

These previous results are used to estimate the counterfactual wage distribution of non-returning migrants, arriving at a similar conclusion to Hanson and Chiquiar (2005). They estimated the human capital distribution of migrants considering only observed characteristics. In both cases it appears that migrants are mainly drawn from the middle part of the distribution of abilities. However, as opposed to Hanson and Chiquiar's estimation, once unobserved heterogeneity is considered, the distribution of abilities of migrants is much more concentrated. This affects all the quintiles of the distribution. We also learn that human capital gains of returning migrants are not a solution to compensate the loss of human capital due to emigration. This is so because, on top of a recent decrease in the size of the flow returning home, it does not appear that returning migrants increase their knowledge much relative to those who stay. We cannot conclude from this paper that emigration is harmful for the country of origin. While it is true that human capital losses appear to be significant and not recovered, remittances may increase the welfare or even the productivity of those who are left behind. A more general approach is required in order to incorporate the effect of remittances into the aggregate human capital of those left behind.

APPENDIX A: Model with observed characteristics

Assume that there is one observed characteristic *X* that is paid differently in each location:

$$\mu^{Mex}(X) = X$$

$$\mu^{US}(X) = \alpha X$$

If *X* is highly correlated with v the model in section 2 applies and the probability of migration would go the other way around of the probability of returning given that someone migrated. We are going to analyse the other extreme case where there is no correlation between the two.

Model with human capital gains

Proposition 1 shows that normality of v ensures that migration and returning migration present an opposite pattern respect to any characteristic. When we change one characteristic, the cutoff for migration and returning vary at the same type. Because of normality, the proportion of returning migrants varies less than the proportion of migrants.

PROPOSITION 1: The probability of migration increases with x if (1-α)<0 , and decreases with x if (1-α)>0 . On the other hand, under some additional mild conditions of the distribution of ability, the probability of returning follows the opposite direction.

PROOF: Let's specify two cases depending on the sign of ρ:

A) (1-ρ)>0

I am going to assume that v has mean equal to 0 and variance equal to 1, although this is not relevant for the proposition. The probability of migration in this case is:

$$P(mig \mid x) = 1 - \Phi\left(\frac{(1-\alpha)X + k + \dfrac{M + R - k}{\pi}}{1-\rho} \right)$$

The derivative then is:

$$\frac{\partial P(mig \mid x)}{\partial X} = -\frac{1-\alpha}{1-\rho}\phi\left(\frac{(1-\alpha)X + k + \frac{M+R-k}{\pi}}{1-\rho}\right)$$

$$sign\left(\frac{\partial P(mig \mid x)}{\partial X}\right) \neq sign(1-\alpha)$$

The probability of returning is:

$$P(ret \mid x, mig) = \frac{\Phi\left(\frac{(1-\alpha)X + k - \frac{R}{1-\pi}}{1-\rho}\right) - \Phi\left(\frac{(1-\alpha)X + k + \frac{M+R-k}{\pi}}{1-\rho}\right)}{1-\Phi\left(\frac{(1-\alpha)X + k + \frac{M+R-k}{\pi}}{1-\rho}\right)} =$$

$$P(ret \mid x, mig) = \frac{\Phi(d_2) - \Phi(d_1)}{1-\Phi(d_1)}$$

Notice that $d_2 > d_1$. The derivative is:

$$\frac{\partial P(ret \mid x, mig)}{\partial X} = \frac{1-\alpha}{1-\rho}\frac{1}{P(mig \mid x)^2}\left[\phi(d_2)(1-\Phi(d_1)) - \phi(d_1)(1-\Phi(d_2))\right]$$

We know by normality that $E(z \mid z > d) = \frac{\phi(d)}{1-\Phi(d)}$, and $0 < \frac{\partial E(z \mid z > d)}{\partial d} < 1$

[see Heckman and Honore (1990)]. Since $d_2 > d_1$ the term in brackets is positive since:

$$\phi(d_2)(1-\Phi(d_1)) - \phi(d_1)(1-\Phi(d_2)) > 0$$

$$\frac{\phi(d_2)}{1-\Phi(d_2)} > \frac{\phi(d_1)}{1-\Phi(d_1)}$$

Therefore,

$$sign\left(\frac{\partial P(ret \mid x, mig)}{\partial X}\right) = sign(1-\alpha)$$

The same reasoning applies to the case (1-ρ)<0.

Therefore, the model of human capital gains gives a simple testable implication that is the probability of migration and the probability of coming back respond differently to changes in characteristics. Let's now see how the addition of uncertainty changes this result.

Model with human capital gains and uncertainty

Proposition 2 shows that there is not enough with normality of v and ε to ensure the same result. However, it is likely that even with uncertainty the same pattern that was observed before will appear. Let's get the intuition with one graph. Figure 16 is the positive selection case in terms of unobserved characteristics. The straight lines represent the cutoffs for a person with characteristic x_1. Let's assume that returns to this particular characteristic in the United States are lower than the returns to the same particular characteristic in México. In this case, a person with a higher x, let's say x_2, will migrate with a lower likelihood, and this is represented in the graph with a right movement of the cuttoff. On the other hand the cutoff of returning migration moves up, which makes every person with characteristics x_2 more willing to come back home. In principle, it appears that the probability of returning given that someone migrated should increase; however, analytically the answer is not that simple. The reason is that all the levels of v that migrate in both situations present a higher probability of coming back with characteristics x_2, but there are some v with a very high probability of coming back with x_1 that with characteristics x_2 decide to stay. In order to get the same result we require this lost to be compensated by the first increase in the probability of coming back. Technically this will depend on the shape of the distributions ε and v.

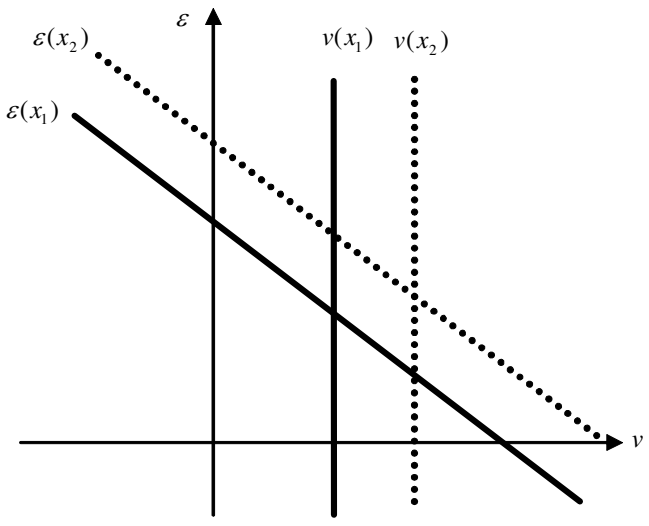


FIGURE 16

PROPOSITION 2: The probability of migration increases with x if (1-α)<0, and decreases with x if (1-α)>0 . On the other hand, the sign of the probability of returning is uncertain.

PROOF: Let's specify two cases depending on the sign of ρ:

A) (1-ρ)>0

All the arguments for the probability of migration apply the same way. Instead, the probability of returning changed a bit. Let's define ε^*

$$\varepsilon^*(v) < (1-\alpha)X + k - \frac{R}{1-\pi} - (1-\rho)v$$

$$P(ret \mid x, mig) = \frac{\int_{d_1}^{\infty} \Phi\left(\varepsilon^*(v)\right)\phi(v)dv}{1-\Phi\left(d_1\right)} =$$

The derivative of the probability of returning is

$$\frac{\partial P(ret \mid x, mig)}{\partial X} = -\frac{\frac{1-\alpha}{1-\rho}(1-\Phi(d_1))\Phi(\varepsilon^*(d_1))\phi\left(d_1\right)}{\left[1-\Phi(d_1)\right]^2} +$$

$$+\frac{(1-\Phi(d_1))\int_{d_1}^{\infty}(1-\alpha)\phi(\varepsilon^*(v))\phi(v)dv}{\left[1-\Phi(d_1)\right]^2} +$$

$$+\frac{\frac{1-\alpha}{1-\rho}\phi(d_1)\int_{d_1}^{\infty}\Phi(\varepsilon^*(v))\phi(v)dv}{\left[1-\Phi(d_1)\right]^2}$$

The first term has a negative sign whereas the other two have a positive one.

However if the distribution of ε is fairly smooth we expect that the loss of migrants with high probabilities of coming back will be compensated by the increase in the probability of returning of all the other migrants. That is the reason why even in this case, we expect the probability of migration and the probability of returning back varying differently with a particular characteristic.

APPENDIX B. Description of the data

10% Mexican Census 2000

The National Institute of Economy, Geography and Computing (the acronym INEGI in Spanish) is the public agency conducting the census in Mexico. The data collection occurred between the 7th and the 18th of February. INEGI used a short questionnaire for the whole population and a long one given to a sample of about 2.2 million housing units. The information was collected by a direct interview in every housing unit. It was accepted that the information given by someone older than 15 whose usual residence was the one in which he was being interviewed knew data from every member of the housing unit. It is a de jure census[25].

There are three levels of interest: individuals, households and housing units. The long questionnaire provides information on socio-demographic characteristics, international migration and household facilities.

Regarding international migration, the reporter must provide information about previous members of the household who migrated abroad between January 1995 and the time of the survey. The question is: "How many people since January 1995 left the country to live abroad?". Additional information regarding those migrants includes gender, age when they migrated, origin and destination, year and month of migration and year and month of returning if they already came back.

In order to get further information about returning migrants, I will need to match migrant histories with that information in the individual file. I know the household of departure, gender and the age when they migrate. Then, I only need to find in that particular household an individual with the same gender and age of the migrant in the year 2000. There are some possible errors. If someone reports to have left in 2000 and in February he is 31 years old, it is possible that the person was 30 at the moment of departure. I kept both possible situations in order to maximise the number of matches. If I find two individuals within the household with the same gender, and with ages differing less than two years I consider the oldest as the migrant. I do the same for the case in which there are two equivalent migrant stories but only one person of the household seems to have those characteristics.

IPUMS 5% US census 2000

The US Census Bureau is the agency conducting census in the US. I use the information from the 5% Public sample. The data collection occurred in April 2000. The Census Bureau also used a short questionnaire for the whole population and a long one given to approximately 17 percent of all housing units (approximately 20 million housing units). The majority of questionnaires were mailed out, and they were mailed back to the Census Bureau before the end of the same month[26]. It is also a de jure census.

---

25. Individuals are considered as members of their usual household, as opposed to being considered members of the household where they are found at the moment of the survey (de facto).

26. 65% of the questionnaires were mailed back by April 19th. The process to reduce problems related to undercoverage was extended until the end of August.

There are two units of interest: individuals and households. The long questionnaire gathers some important information about socio-demographic characteristics and household facilities.

There are two questions that determine whether a resident is an immigrant: the place of birth, and when the person came to live in the United States.

The American census does not ask anything related to legal status. Therefore, it should capture people who are immigrant aliens (those with a valid visa of residence), temporary visitors (students regardless whether they are in an exchange program and temporary workers) and undocumented. Regarding temporary visitors, the census is not expected to capture tourists, very short term visitors and commuters because individuals who live or stay at another country most of the time do not fill out the questionnaire. On the other hand, some illegal immigrants may have a high incentive to hide their situation in the US. This may generate underreporting of immigrants. In later sections, I will discuss underreporting and the possible biases that this phenomenon generates.

APPENDIX C: Probits for the decision of leaving and coming back

Panel A (Individual characteristics)

	Migration Migrant Cohort 95-00	Return >3 months
Age	-0,0030	0,0161
	(0,000017)**	(0,000182)**
Agesq	0,0000	-0,0002
	(0,000000)**	(0,000003)**
Education	-0,0165	0,1442
1-8 courses	(0,000124)**	(0,001694)**
	-0,0227	0,1775
9-11 courses	(0,000092)**	(0,002367)**
	0,0038	0,0034
12 courses	(0,000158)**	-0,0017
	-0,0153	0,0615
>12 courses	(0,000101)**	(0,002345)**
	0,0038	0,0306
Marriage	(0,000069)**	(0,000703)**
	0,5596	-0,0902
Other HH migration in	(0,001014)**	(0,000629)**
Observations	22.800.000	917.877

Standard errors in parentheses
* significant at 5%; ** significant at 1%

Panel B (Household characteristics)

	Migration Migrant Cohort 95-00	Return >3 months
Size HH	0,0011	-0,0427
	(0,000085)**	(0,000839)**
Children HH	-0,0015	0,0497
	(0,000097)**	(0,000931)**
Eldery HH	0,0005	0,0227
	(0,000244)*	(0,001964)**
People working in HH	0,0080	-0,0469
	(0,000090)**	(0,000876)**
Students in HH	0,0026	-0,0102
	(0,000135)**	(0,001178)**
Other income apart from	0,0000	0,0000
wages and remittances	(8,49e-08)**	(7,86e-07)
	0,1905	-0,1184
Family Migrating 1995	(0,001125)**	(0,001626)**
Region	Yes	Yes
Size of the Village	Yes	Yes
Observations	22.800.000	917.877
Standard errors in parentheses		
* significant at 5%; ** significant at 1%		

Source - Panel A: Mexican and American censuses in 2000. Sample restricted to Males between 15 and 65 in 1995 who have positive potential experience and Panel B: Mexican census in 2000. Sample restricted to Males between 15 and 65 in 1995 who have positive potential experience. Household characteristics include the migrant. If he is <15 he is considered a student. If he is >65 he is considered elderly. Otherwise the person is considered to be working. Probit of migration: 1 is migrant between 1995 and 1999 regardless the returning status. Probit of coming back: 1 is a migrant who left between 1995 and 1999 and came back after at least 3 months abroad.

REFERENCES

BEENSTOCK, M. (1996). "Failure to Absorb: Remigration by Immigrants into Israel", *International Migration Review,* Vol. 30, No. 4. Winter, pp. 950-978.

BORJAS, G.(1987). "Self-Selection and the Earnings of Immigrants", *The American Economic Review,* Vol. 77, No. 4., pp. 531-553.

BORJAS, G., and B. BRATSBERG (1996). "Who leaves? The Outmigration of the foreign born", *The Review of economics and statistics,* Vol. 78, issue 1, pp. 165-176.

BORJAS, G., R. FREEMAN and K. LANG (1991). "Undocumented Mexican-Born Workers in the United States: How Many, How Permanent?" in J. Aboud and R. Freeman, *Immigration, Trade, and the Labor market,* NBER, Chicago, The University of Chicago Press.

BUSTAMANTE, J., G. JASSO, J. TAYLOR and P. TRIGUEROS (1997). "The selectivity of International labor migration and characteristics of Mexico-to-US Migrants", in Mexico-United States Binational Migration Study *Migration between Mexico and the United States,* Mexican Ministry of Foreign Affairs and US Comission on Immigration Reform, Vol. 2.

CHISWICK, B. (1978). "The Effect of Americanization on the Earnings of Foreign-born Men", *The Journal of Political Economy,* Vol. 86, No. 5, pp. 897-921.

DAVANZO, J. (1976). "Differences between Return and Nonreturn Migration: An Econometric Analysis", *International Migration Review,* Vol. 10, No. 1, pp. 13-27.

DELEIRE, T., B. JACOB, A. LACUESTA and R. LALONDE (2005). *The Americanization of Immigrant Children from Mexico,* Working Paper.

DINARDO, J., M. FORTIN and T. LEMIEUX (1996). "Labor market institutions and the distribution of wages, 1973-1992: a semiparametric approach", *Econometrica,* Vol. 64, No. 5, pp. 1001-1044.

HANSON, G., and D. CHIQUIAR (2005). "International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States", *Journal of Political Economy*, Vol. 113, No. 2.

HECKMAN, J., and B. HONORE (1990). "The Empirical Content of the Roy Model", *Econometrica*, Vol. 58, No. 5. pp. 1121-1149.

HEER, D., and J. PASSEL (1987). "Comparison of two different methods for computing the Number of Undocumented Mexican Adults in Los Angeles County", *International Migration Review,* 21, pp. 1446-1473.

LACUESTA, A. (2005). *Emigration and human capital: who leaves, who comes back and what difference does it make?,* Thesis (Ph. D.) University of Chicago, Dept of Economics.

LINDSTROM, D. (1996). "Economic Opportunity in Mexico and Return Migration from the United States", *Demography*, Vol. 33, No. 3, pp. 357-374.

LONG, L. H., and K. A. HANSEN (1975). "Trends in Return Migration to the South", *Demography*, Vol. 12, No. 4., pp. 601-614.

MASSEY, D., and K. ESPINOSA (1997). "What's Driving Mexico-US Migration? A Theoretical, Empirical, and Policy Analysis", *The American Journal of Sociology*, Vol. 102, No. 4., pp. 939-999.

MASSEY, D., and R. ZENTENO (2000). "A Validation of the Ethnosurvey: The case of Mexico-US Migration", *International Migration Review,* Vol. 34, pp. 766-793.

BANCO DE ESPAÑA PUBLICATIONS

WORKING PAPERS[1]

0527 LUIS J. ÁLVAREZ, PABLO BURRIEL AND IGNACIO HERNANDO: Price setting behaviour in Spain: evidence from micro PPI data.
0528 EMMANUEL DHYNE, LUIS J. ÁLVAREZ, HERVÉ LE BIHAN, GIOVANNI VERONESE, DANIEL DIAS, JOHANNES HOFFMANN, NICOLE JONKER, PATRICK LÜNNEMANN, FABIO RUMLER AND JOUKO VILMUNEN: Price setting in the euro area: some stylized facts from individual consumer price data.
0529 TERESA SASTRE AND JOSÉ LUIS FERNÁNDEZ-SÁNCHEZ: Un modelo empírico de las decisiones de gasto de las familias españolas.
0530 ALFREDO MARTÍN-OLIVER, VICENTE SALAS-FUMÁS AND JESÚS SAURINA: A test of the law of one price in retail banking.
0531 GABRIEL JIMÉNEZ AND JESÚS SAURINA: Credit cycles, credit risk, and prudential regulation.
0532 BEATRIZ DE-BLAS-PÉREZ: Exchange rate dynamics in economies with portfolio rigidities.
0533 ÓSCAR J. ARCE: Reflections on fiscalist divergent price-paths.
0534 M.ª DE LOS LLANOS MATEA AND MIGUEL PÉREZ: Differences in changes in fresh food prices by type of establishment. (The Spanish original of this publication has the same number.)
0535 JOSÉ MANUEL MARQUÉS, FERNANDO NIETO AND ANA DEL RÍO: Una aproximación a los determinantes de la financiación de las sociedades no financieras en España.
0536 S. FABIANI, M. DRUANT, I. HERNANDO, C. KWAPIL, B. LANDAU, C. LOUPIAS, F. MARTINS, T. MATHÄ, R. SABBATINI, H. STAHL AND A. STOKMAN: The pricing behaviour of firms in the euro area: new survey evidence.
0537 LUIS J. ÁLVAREZ AND I. HERNANDO: The price setting behaviour of Spanish firms: evidence from survey data.
0538 JOSÉ MANUEL CAMPA, LINDA S. GOLDBERG AND JOSÉ M. GONZÁLEZ-MÍNGUEZ: Exchange-rate pass-through to import prices in the euro area.
0539 RAQUEL LAGO-GONZÁLEZ AND VICENTE SALAS-FUMÁS: Market power and bank interest rate adjustments.
0540 FERNANDO RESTOY AND ROSA RODRÍGUEZ: Can fundamentals explain cross-country correlations of asset returns?
0541 FRANCISCO ALONSO AND ROBERTO BLANCO: Is the volatility of the EONIA transmitted to longer-term euro money market interest rates?
0542 LUIS J. ÁLVAREZ, EMMANUEL DHYNE, MARCO M. HOEBERICHTS, CLAUDIA KWAPIL, HERVÉ LE BIHAN, PATRICK LÜNNEMANN, FERNANDO MARTINS, ROBERTO SABBATINI, HARALD STAHL, PHILIP VERMEULEN AND JOUKO VILMUNEN: Sticky prices in the euro area: a summary of new micro evidence.
0601 ARTURO GALINDO, ALEJANDRO IZQUIERDO AND JOSÉ MANUEL MONTERO: Real exchange rates, dollarization and industrial employment in Latin America.
0602 JUAN A. ROJAS AND CARLOS URRUTIA: Social security reform with uninsurable income risk and endogenous borrowing constraints.
0603 CRISTINA BARCELÓ: Housing tenure and labour mobility: a comparison across European countries.
0604 FRANCISCO DE CASTRO AND PABLO HERNÁNDEZ DE COS: The economic effects of exogenous fiscal shocks in Spain: a SVAR approach.
0605 RICARDO GIMENO AND CARMEN MARTÍNEZ-CARRASCAL: The interaction between house prices and loans for house purchase. The Spanish case.
0606 JAVIER DELGADO, VICENTE SALAS AND JESÚS SAURINA: The joint size and ownership specialization in banks' lending.
0607 ÓSCAR J. ARCE: Speculative hyperinflations: When can we rule them out?
0608 PALOMA LÓPEZ-GARCÍA AND SERGIO PUENTE: Business demography in Spain: determinants of firm survival.
0609 JUAN AYUSO AND FERNANDO RESTOY: House prices and rents in Spain: Does the discount factor matter?
0610 ÓSCAR J. ARCE AND J. DAVID LÓPEZ-SALIDO: House prices, rents, and interest rates under collateral constraints.
0611 ENRIQUE ALBEROLA AND JOSÉ MANUEL MONTERO: Debt sustainability and procyclical fiscal policies in Latin America.

---

1. Previously published Working Papers are listed in the Banco de España publications calalogue.

0612 GABRIEL JIMÉNEZ, VICENTE SALAS AND JESÚS SAURINA: Credit market competition, collateral and firms' finance.

0613 ÁNGEL GAVILÁN: Wage inequality, segregation by skill and the price of capital in an assignment model.

0614 DANIEL PÉREZ, VICENTE SALAS AND JESÚS SAURINA: Earnings and capital management in alternative loan loss provision regulatory regimes.

0615 MARIO IZQUIERDO AND AITOR LACUESTA: Wage inequality in Spain: Recent developments.

0616 K. C. FUNG, ALICIA GARCÍA-HERRERO, HITOMI IIZAKA AND ALAN SUI: Hard or soft? Institutional reforms and infraestructure spending as determinants of foreign direct investment in China.

0617 JAVIER DÍAZ-CASSOU, ALICIA GARCÍA-HERRERO AND LUIS MOLINA: What kind of capital flows does the IMF catalyze and when?

0618 SERGIO PUENTE: Dynamic stability in repeated games.

0619 FEDERICO RAVENNA: Vector autoregressions and reduced form representations of DSGE models.

0620 AITOR LACUESTA: Emigration and human capital: Who leaves, who comes back and what difference does it make?

BANCO DE ESPAÑA

Unidad de Publicaciones
Alcalá, 522; 28027 Madrid
Telephone +34 91 338 6363. Fax +34 91 338 6488
e-mail: Publicaciones@bde.es
www.bde.es