Self selection of employed rural-metropolitan migrants in Paraguay. The role of education

Carlos Villalobos Barría*

Complete draft, February, 2010

Abstract

Given the different income inequalities between rural and metropolitan areas in Paraguay, the standard Roy model predicts the possibility that a selection bias of rural metropolitan migrants exists. Based on an extended Roy model, which allows for correlation between labour market characteristics and moving costs, a switching regression model with endogenous switching was used to evaluate if employed migrants from rural to metropolitan areas are somehow selected based on unobserved characteristics. Regarding unobserved characteristics, evidence supports that rural metropolitan migrants are negatively selected only in 1997. In 2005, they appear to be a random sample of the rural labour force. Regarding observed characteristics, such as education, I find that migrants are negatively selected.

Additionally, considerable evidence is provided to reveal the dual impact of education on rural metropolitan migration probability. On the one hand, individuals own education discourages migration; on the other, the level of household education encourages migration by increasing the degree of freedom of the potential migrant. Ignoring this dual impact of education yields to obtain biased impacts of education on migration, and therefore the unobservable effect on migration probability, could be misleading.

This paper offers the first attempt to estimate a structural probability model for employed rural metropolitan migrants in Paraguay and finds that migration is not only driven by market signals such as wage differentials or unemployment rates. Gender, marital status and the number of children per household are always relevant and significant determinants of rural metropolitan migration of workers in a structural form in Paraguay.

***JEL* Classification**: R23, J31, J61 **Keywords**: migration, wages, self selection

*Development economics research group at the University of Göttingen (cvillal@uni-goettingen.de). I would like to thank Stephan Klasen for his helpful and useful suggestions and to the participants of the third ECINEQ annual meeting held in Buenos Aires, Argentina, 21-23 July 2009, and in particular François Bourguignon and François Maniquet for helpful discussions and comments. I am also grateful to the participants of The International Society for Comparative Economic Studies, ISCES, 9th Arnoldshain Seminar 2009, held in Montevideo, Uruguay and the participants of the European Development Research Network, EUDN, 8th Workshop for Ph.D Students August 2009 in Örebro, Sweden for the useful discussion.

1 Introduction

Migration movements in Paraguay are considerable. Rural, urban and metropolitan areas work as both source and destination areas. As a result of human capital levels being considerably lower in rural areas than in metropolitan ones, my particular interest focuses on studying rural to metropolitan migration and how rural migrants integrate into metropolitan labour markets. Migration has consequences for households and regions and may work as a mechanism to counterbalance relative resource scarcities over regions. Individual migration decisions respond to economic opportunities as migrants seek higher returns to their attributes [1]. Migration in Paraguay has been almost overlooked by empirical researchers. Nevertheless, there are significant works that give some highlights on this issue. Richards (1990) analyzes the emigration to Buenos Aires from a gender perspective, Rivarola (1967) analyzes migration in Paraguay from a sociological point of view. Molinas (1999) studies the internal migration using the 1996 household survey. Galeano (1997) exploits census data for a descriptive analysis. Richards (1996) focuses on migration from Asuncion to Buenos Aires and Sosa (1996) analyses the role of gender on migration.

The decision to migrate is based on the perspective that the destination regions will be able to provide better levels of welfare to the migrants, in relation to what they would experience in their place of origin. There are numerous factors that can explain the decision to migrate to a certain area. Among these are the relative indices for standards of living, unemployment, household characteristics, age, education attainment of the household head, etc. However, there are some non-observable characteristics of the individuals that influence the decision to migrate and also their level of earnings at their destination. The composition of the migrant population is key to evaluating the consequences of migration. This composition concerns human capital characteristics and non-observable abilities. For this reason, the study of self-selection of migrants, taking the non-observed ability into account, plays an important role in understanding the causes of the rural metropolitan migration and their effects on labour markets, demand for public goods, public expenditure, investment, poverty and overall prospects of economic development.

The motivation for migration studies varies in relation to the scope of view. In particular, flows of migrants workers from developing to developed countries are a subject of study because of the potential impact that a negative selection of migrants could have on the developed economies, given their magnitudes[2] and of course the impact on the development prospects for the underdeveloped economies of origin (Krugman, 1991 and Coniglio, 2003). Rural metropolitan migration

[1] Returns can be monetary or non monetary. For example the role of amenities has been discussed in Roback (1982), Hoehn, Berger and Blomquist (1987).

[2] See Zlotnik (1998) and Stalker (2000)

of workers between two regions of a country is of great interest because, given the relative lower cost of migration, a negative selection of migrants is possible and therefore, the study of the rural metropolitan migration of workers is crucial to understand the underlying determinant of a successful labour assimilation in the destination areas. This is also very important for policy makers to prevent dysfunctional migration flows that are translated into marginalization and poverty around metropolitan areas.

Two types of models have been developed in an attempt to explain migration within countries: (a) models based on the Harris-Todaro model (1970) which states that the migration decision is based on expected income differentials between rural and urban areas and not on wage differentials, implying that rural-urban migration in a context of high urban unemployment can be economically rational if expected urban income exceeds expected rural income. Fields (1975) extends the Harris-Todaro model allowing preferential hiring by educational level and other extensions[3]. (b) Models based on the human capital theory (Becker 1964, Schultz 1961) and the maximization of utility. For example Greenwood (1997) states that migration occurs if the place that maximizes net utility differs from the current location. The human capital models on migration states that variables concerning the life-cycle of individuals are also determinants of migration. In this context the individual preferences and endowments change over time in relation to his life-cycle stadium[4].

In summary, individuals are assumed to move if the returns (expected) to migration are greater than the sum of all opportunity costs and the moving costs (Borjas, 1987; Jacob Mincer, 1978). Given this benefit-cost analysis, only a determined group of individuals and households will have an economic incentive to emigrate. In this context, positive selection means that, relative to the home-region population (labour force), movers are of above average quality. While common wisdom maintains that movers are the most motivated individuals of the home-region population (labour force), it may be possible that individuals (workers) from the bottom tail of the quality distribution are more likely to migrate due to the economic incentives.

Borjas (1987) used the Roy (1951) model to derive an economic model of selection based on unobserved characteristics. In this model, the immigrants are negative selected if they have below average wages and productivities given their total characteristics (observed and unobserved). Conversely, positive selected immigrants have above average earnings and productivities. Borjas shows that there is no relationship between the selection process generated by unobserved characteristics and the selection process generated by observed characteristics (i.e education). In

[3]Four extensions of the Harris-Todaro model are made, including allowances for more generalized job-search behaviour, an urban traditional sector, preferential hiring by educational level, and labour turnover considerations.

[4]: Clark and Hunter (1992) show that life-cycle position may explain destination choices, where particular characteristics of potential destinations meet the requirements of different population groups.

concordance with the kernel of the Harris-Todaro model, highly educated individuals are more able to access good quality information on labour markets in potential destinations and therefore highly educated people are more likely to get a job than unskilled workers if the correlation between labour markets in both regions is strong enough. However, as Borjas (1988) states, "it is completely possible for the most educated persons to migrate to the U.S (i.e., positive selection in education), but for these persons to be the least productive persons in the population of highly educated persons (i.e., negative selection in unobserved characteristics)." There are consequently two dimensions of "quality", and therefore generalizations based only on observed characteristics are misleading, because of much more than a half of the variance in wage rates or weekly earnings is explained by unobserved characteristics.

The standard Roy model predicts that migrants will be negative selected if the inequality in the distribution of incomes is lower in the destination than in the origin and if the correlation between wages in both regions is strong enough. Since the Gini-inequality in earnings in Paraguay is higher for rural areas than for metropolitan areas[5], the Roy model predicts that rural metropolitan migrants are negative selected concerning unobserved ability.[6] However a generalization of the Roy model that relaxes the assumption of constant moving costs by allowing correlation between non-observed abilities and moving costs, shows that the type of selection may change in either direction (Borjas, 1987). Highly educated individuals are more likely to access good quality information on labour markets in potential destinations and therefore make rational decisions concerning migration. For this reason, highly educated people are more likely to get a job than unskilled workers if the correlation between labour markets in both regions is strong enough. However, in countries with a low human capital accumulation in rural areas it is highly plausible that educated workers are less able to move because of the high possibility of finding well paid work within their regions of origin, and therefore negative selection regarding education levels can also occur (Richards, 1996).

The empirical approach of this paper consists of a switching regression model presented by Goldfeld and Quandt (1973) with endogenous switching (Maddala, 1983) which can be properly used as counterpart from an extended Roy model in the migration context.

This paper contributes through the provision of a structural form for rural metropolitan migration

[5]The Gini coefficients are calculated based on the hourly wages for all occupations reported by employed labour force participants. The fact that wage inequality is higher in rural areas than in metropolitan areas, holds for both analyzed cohorts.

[6]Recent literature finds evidence that migrants are in fact positive selected. Chiswick (2000) states "the analysis indicates a tendency toward the favourable self-selection (supply) of migrants for labour market success on the basis of a higher level of ability broadly defined. The favourable selectivity is more intense: the greater the out of pocket (direct) costs of migration and return migration, the greater the effect of ability on lowering the costs of migration, and the smaller the wage differences by skill (...)".

which is derived from a powerful theoretical framework[7]. The results of these structural probability models are robust through years and aggregation levels of the destination and supports the main concepts of the Harris Todaro model. Expected wage differentials, unemployment rates and well controlled household characteristics play an important role regarding migration. Nevertheless, the great contribution of this paper consists in dismembering and highlighting the role of education on migration. On the one hand, individual's education discourages migration, while on the other; the level of household education encourages migration by increasing the degree of freedom of the potential migrant. Ignoring this dual impact of education results in biased impacts of education on migration and consequently the effect of unobservable on migration probability could be misleading.

This paper proceeds as follows: Section 2 reveals the data and sample characteristics. Section 3 offers a descriptive view concerning the migration flows focusing on rural metropolitan migration patterns addressing issues relative to the role of education level of the family, ex-post selection of employed migrants and earning inequalities. The theoretical framework based on the extended Roy model is presented in Section 4, followed by the results for the reduced forms of the probability to move, the Heckit wage regressions and the structural probability models in Section 5, and a conclusion in Section 6.

[7] As far as the author is aware, this is the first attempt to explain migration in Paraguay in a structural way.

2 The Data

For this study, the official data was provided by the Dirección General de Estadísticas, Encuestas y Censos (DGEEC) de Paraguay. The Encuesta Permanente de Hogares 1992 and 2005 (*EPH 1992, 2005*) and the Encuesta Integral de Hogares 1997/98 and 2000/01 (*EIH 1997/98, 2000/01*) were also used. The studied cohorts (*EIH 1997/98* and *EPH 2005*) are a two-step sampling living standard survey which offers each household the same probability of being sampled. Only the departments of Boquerón and Alto Paraguay have not been represented in the survey.[8]

The *EIH 1997/98* consists of 20,064 observations corresponding to 4,353 surveyed households. The *EPH 2005* comprises 19,579 observations from 4,464 households. Both surveys are representative at country, urban, rural, strata and main department levels.[9] District aggregated lagged values of the mean unemployment rates and mean household income have been used taking into consideration that the decision to migrate was made some time before the studied year. These were obtained from an available previous survey with similar characteristics. Lagged values for 2005 and for 1997/98 were obtained from *EIH 2000/2001* and from *EPH 1992*, respectively. .

The *EIH 2000/2001* comprises 8,131 households and the same representativeness as the principal surveys. The EPH 1992 is representative only at metropolitan and non-metropolitan levels. Therefore, the lagged variables from 1998 used in this study, could be potentially biased. Nevertheless, a detailed data revision was carried out in order to detect inconsistencies that could bias the estimates.[10]

In this paper, a migrant is defined as an individual who has moved to a different district within the past five years, excluding migrants from foreign countries[11]. The study focuses on the rural metropolitan migration pattern of employed individuals. This pattern implies that the studied sample is composed of all individuals who report to be domiciled in rural areas five years ago and according to that, two categories of individuals have been distinguished. *Movers* are therefore the

[8] Boqueron and Alto Paraguay were excluded due to the low concentration of population (less than 2% of the national population) and the large area to be covered implying an extremely high operational cost. For more detailed information, see http://www.dgeec.gov.py/MECOVI/.

[9] Representative for San Pedro, Itapúa, Caaguazú, Alto Paraná and Central departments.

[10] No inconsistencies were founded. Income rankings among departments held respect to other surveys and are consistent with other sources of information concerning relative development indicators.

[11] Due to the requirements of our econometric estimation concerning a minimum number of observations for migrants, it was not possible to define a migrant as an individual who has moved to a different district within the past years (*recent migrant*). Consequently, workers that have migrated and later returned to rural areas during this period are not considered as migrants. This variable could also capture *"successful migration"*, rather than *"whole migration"*, and therefore our estimates on self-selectivity could be potentially more biased towards positive selection regarding the ability distribution.

employed individuals who moved from rural to metropolitan areas any time within the last five years, while *stayers* are the employed individuals who continue to live in rural areas.

Table 1: Sample size 1997, employed rural origin individuals

	Rural Stayers	Movers[1]	Movers[2]	Total
Observations	1857	84	168	1981
Population Size	413319	25007	49896	463215

[1]:to metropolitan area,[2]:to metropolitan and urban Areas

Source: own elaboration, based on EIH 1997/98

Table 2: Sample size 2005, employed rural origin individuals

	Rural Stayers	Movers[1]	Movers[2]	Total
Observations	1140	45	95	1235
Population Size	312227	14303	26205	338432

[1]:to metropolitan area,[2]:to metropolitan and urban Areas

Source: own elaboration, based on EPH 2005

Tables 1 and 2 show the relevant sample for this study. Although I have focused on employed rural metropolitan migration, the analysis also takes into consideration unemployment and other destination aggregation levels for robustness purposes. On the one hand, ignoring unemployment rates in the analysis may bias the estimates if unemployment has an unequal effect on movers and stayers after controlling for observed characteristics; on the other hand, running the analysis by using the rural to urban/metropolitan migration pattern is to rule out the possibility that the result could be drove driven by the fact that some rural metropolitan migrants are coming from directly neighbouring districts, only a few kilometres away, towards the great agglomeration around Asunción and therefore not necessarily representing a genuine migration process[12].

Tables 3 and 4 show the simple sizes considering the labour force participants. It is noteworthy, that the number of observations increases considerably thus making easier the modelling of the probability specifications.

Table 3: Sample size 1997, rural origin labor force particiants

	Rural Stayers	Movers[1]	Movers[2]	Total
Observations	3029	89	199	3118
Population Size	661407	27680	58447	723907

[1]:to metropolitan area,[2]:to metropolitan and urban Areas

Source: own elaboration, based on EIH 1997/98

[12]The urban/metropolitan destination is spatially more equally distributed than the "only metropolitan" aggregation level used. An additional advantage is that the portion of migrants growth significantly, allowing the study of whether some results are only driven by the relatively few migrants in the sample.

Table 4: Sample size 2005, rural origin labor force participants

	Rural Stayers	Movers[1]	Movers[2]	Total
Observations	2403	49	108	2511
Population Size	670875	16322	31014	701889

[1]:to metropolitan area,[2]:to metropolitan and urban Areas

Source: own elaboration, based on EPH 2005

3 Migration flows in Paraguay

A migrant is defined as an individual who has moved to a different district within the last five years, excluding migrants from foreign countries. The EPH identifies the birthplace of migrants. Analyzing only recent migration and not lifetime migration has the drawback of not allowing the study of their integration at destination. Restricting the sample to a more homogeneous group makes the analysis of the migration decision more meaningful. Asuncion and some districts of the Central department have been defined as the metropolitan area. All districts, which have not been defined as rural areas, belong to urban areas.

Although family members usually migrate together, it is reasonable to attribute the migration decision to the household head. In some circumstances the head migrates first, accumulates income and then moves the rest of the family. In others, the head might even stay apart from his/her family and send remittances back home. Tables 1 and 2 present the migration flows for household heads aged 15 to 65 between rural, urban and metropolitan areas in 1997-98 and 2005, respectively. A location is classified as rural, urban or as metropolitan following the definition used in the EPH[13].

Through the observation of data, it can be seen that the rural-urban/metropolitan migration has a small size and becomes relatively less important in relation to other kinds of migration patterns, such as urban-urban, urban-metro, metro-urban, metro-metro, metro-rural and urban-rural migrations. In spite of the relatively small size of the rural-urban/metropolitan migration, it should represent a radical life-style change and therefore, the study this particular migration pattern continues to be of interest. Descriptive statistics show that rural metropolitan migrants are characterized by being younger, having smaller families and no longer being less educated than non-migrant individuals[14].

[13]Encuesta Permanente de Hogares defines an urban area as a locality formed by ten or more blocks, independently of the number of inhabitants. The metropolitan area comprises Asuncion and the following districts of the *Central department*: C*apiatá, Fernando de la Mora, Guarambaré, Itá, Itaguá, Lambaré, Limpio, Luque, Mariano Roque Alonso, Nueva Italia, Nemby, San Antonio, San Lorenzo, Villa Elisa, Villeta.*

[14]For 1997 and 2005, rural-metro and rural-urban migrants are on average more educated than rural non-migrants. This is an indication of the "positive" (unconditional) selection of this group of migrants regarding years of schooling.

Table 5: Migration flows, by origin and destination, 1997/98 (heads only)

Origin	N° of Migrants	% of destination flow	% of origin flow	% of total flow
Migrants to Rural Areas				
Rural	23,417	49.23	51.30	21.30
Urbano	17,050	35.85	36.37	**15.51**
Metropolitano	7,095	14.92	40.71	6.45
Total	47,562	100.00	43.26	43.26
Migrants to Urban Areas				
Rural	11,267	29.86	24.68	10.25
Urbano	16,129	42.75	34.40	**14.67**
Metropolitano	10,332	27.39	59.29	9.40
Total	37,728	100.00	34.31	34.31
Migrants to Metropolitan Area				
Rural	10,961	44.44	24.01	**9.97**
Urbano	13,706	55.56	29.23	**12.46**
Total	24,667	100.00	17.86	17.86
Migrants Rural, Urban and Metropolitan Areas				
Rural	45,645	**41.51**	100.00	41.51
Urbano	46,885	**42.64**	100.00	42.64
Metropolitano	17,427	15.85	100.00	15.85
Total	109,957	100.00	100.00	100.00

Source: own elaboration, based on EIH 1997/98

In 1997-98, Table 5 shows that the main flow of migrating household heads comes from urban areas (42.64%) with 14.67% moving to other urban areas, 12.46% moving to metropolitan destinations and 15.51% moving to rural areas. The second main flow in 1997/98 comes from rural areas and represents 41.51% of the total migration[15]. The rural metropolitan migration pattern represents 9.97% of the total migration.

In 2005 (Table 6), the main flow of migrating household heads comes from urban areas and represents 45.33% of the total migration. This figure consists of 20.30% moving to rural areas, 12,47% moving to other urban areas and 12.55% moving to metropolitan areas. The rural metropolitan migration pattern represents 5.62% of the whole migration.

In 1997/98, household heads leaving rural areas represented 41.51% of the total migration. In 2005, this component was 35.93%. For this paper, it is of particular interest that the rural movers to urban or metropolitan areas represent 20.22% of the total migration in 1997/98, while repre-

[15]Metropolitan to metropolitan migration pattern corresponds to a district change within Asunción. It is highly likely that this migration pattern does not correspond to a genuine migration process and is therefore not comparable to the rural metropolitan migration pattern, which represents a genuine one, with all their associated consequences. This paper is not interested in studying this kind of migration pattern. Metropolitan to metropolitan migration pattern has been dropped in this table resulting in an increment in the relative importance of the rural metropolitan migration pattern. The full tables, containing all identified and exhaustive migration patterns are presented in the Appendix.

Table 6: Migration flows, by origin and destination, 2005 (heads only)

Origin	N° of Migrants	% of destination flow	% of origin flow	% of total flow
Migrants to Rural Areas				
Rural	23,928	42.40	67.29	24.18
Urbano	20,105	35.62	44.82	**20.32**
Metropolitano	12,403	21.98	66.91	12.53
Total	56,436	100.00	57.03	57.03
Migrants to Urban Areas				
Rural	6,066	24.72	17.06	6.13
Urbano	12,342	50.29	27.51	**12.47**
Metropolitano	6,135	25.00	33.09	6.20
Total	24,543	100.00	24.80	24.80
Migrants to Metropolitan Area				
Rural	5,563	30.94	15.65	**5.62**
Urbano	12,415	69.06	27.67	**12.55**
Total	17,978	100.00	13.18	13.18
Migrants Rural, Urban and Metropolitan Areas				
Rural	35,557	35.93	100.00	35.93
Urbano	44,862	**45.33**	100.00	45.33
Metropolitano	18,538	18.73	100.00	18.73
Total	98,957	100.00	100.00	100.00

Source: own elaboration, based on EPH 2005

senting 11.75% of the total in 2005[16].

Figure 1 shows the hourly log wages[17] at current prices for rural migrants and non-migrants in 1997 and 2005. On average, migrants in both cohorts appear to receive higher wages than their non-migrant counterparts. Otter and Villalobos (2009) present evidence that migration explains higher returns only for low quantiles in the conditional wage distribution[18].

3.1 The role of the education levels of the family

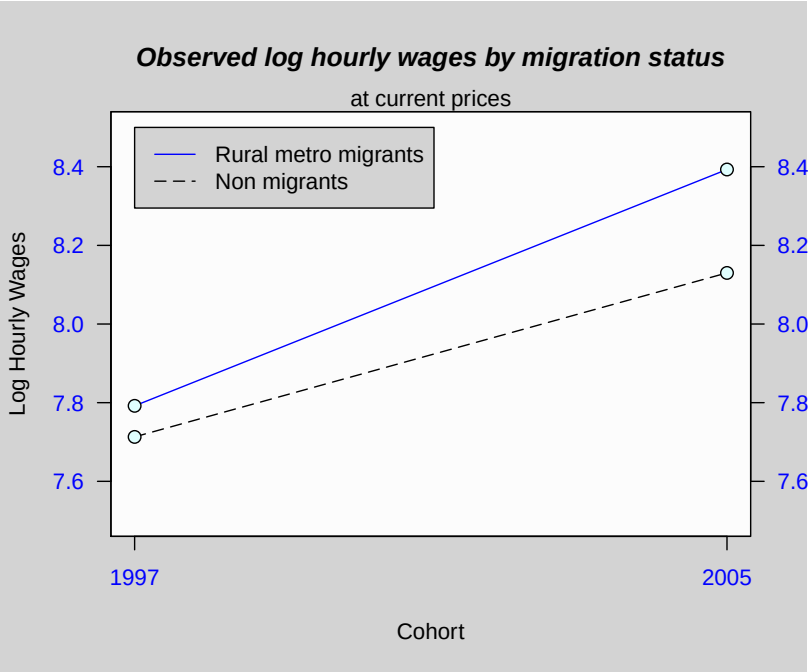
The degree in which the remaining members of the household can generate income is also important information for the potential migrant. Households with lower levels of human capital accumulation are more dependent on the monetary and non-monetary contribution of each member. The cost in case of migration will be somehow proportional to the household welfare dependence of any household member and therefore, an individual would be able to move if he knows that in

[16]The figures presented in the tables were drafted using the sample weights provided in both surveys.

[17]Reported hourly wages for all occupations.

[18]Otter and Villalobos (2009) use quantile wage regressions to determine if migration explains wages. In the heteroscedastic Paraguayan data, they find significant coefficients only for Q (0.1) and Q (0.25) 0.411 and 0.285 at the 10% level respectively.

Figure 1: Observed log hourly wages by migration status



Source: Own elaboration, based on EIH 1997/98 and EPH 2005

case of migration, the household will not fall into poverty[19]. I introduce this cost proposition by using the average family education in years as an indicator for the household capacity of income generation.

It would appear that this structural variable has never been used in empirical studies. For instance, and for a variety of micro and macro model specifications to test for self-selection of migrants, Yashiv (2004), Brücker and Jahn (2008), Cobb and Clark(1993), Cohen and Haberfeld (2007), Brücker and Defort (2006), Jensen et al (2006), Parrado and Cerrutti (2003) and Abramitzky (2008) used only individual education variables in their models. Fafchamps and Shilpi (2008) additionally use the father's education level in the migration selection equation.

3.2 Distribution of migrants among wages and education levels

In an attempt to better understand the selection of migrants, a comparison was made between the wages, individual education and average household education of employed rural metropolitan migrants and the levels of their respective stayer counterparts. By doing so, it was possible to obtain some information to help identify those characteristics that can differ in levels between the

[19]The term "fall into poverty" is used as a particular example from the more general concept "a considerable reduction of the household welfare".

two sub-samples.

In general, if self-selection of migrants exists, it can be attributed to the sum of the selection process regarding observed and unobserved characteristics. As wages are explained in the same way by observed and unobserved characteristics, the migrant sub-sample is divided among high, middle and low wages to explore whether rural metropolitan migrants come principally from the bottom, middle or upper tail of the wages distribution[20]. Tables 7-10 present these figures including the average years of education at individual and household levels[21].

Table 7: Rural-Metro Migrant strucure and average education levels for individuals and their households, 1997/98

	Rural-Metro migrants		
Wage group	Share(%)	Years of education	Years of household education
Low	13.98	5.63	5.95
Middle	73.17	6.44	6.87
High	12.85	8.10	8.27

All average of education expresed in years
Source:own elaboration, based on EIH 1997/98

Table 8: Rural Stayer strucure and average education levels for individuals and their households, 1997/98

	Rural stayers		
Wage group	Share(%)	Years of education	Average household education
Low	56.10	4.22	3.37
Middle	27.67	5.49	4.19
High	16.23	7.35	5.01

All averages of education are expresed in years
Source:own elaboration, based on EIH 1997/98

Table 9: Rural-Metro Migrant strucure and average education levels for individuals and their households, 2005

	Rural-Metro migrants		
Wage group	Share (%)	Years of education	Years of household education
Low	4.86	7.50	7.50
Middle	53.95	8.32	8.50
High	41.19	9.63	9.56

All average of education expresed in years
Source:own elaboration, based on EPH 2005

[20]the wage distribution has been divided into three groups; the bottom tail of the distribution (0%-33,3%), the middle tail of the distribution (33.4%-66.6%) and the upper tail of the distribution which represent the higher 33.3% of wages. Workers who are in the bottom of the distribution are more likely to be poor. Therefore, the distribution of migrants among these income groups is a similar way to analyze how the total selection process concerns observed and unobserved characteristics as a whole. If migrants come from the bottom tail of the wage distribution, it would be evidence of a negative selection process. Contrary to this, if migrants are more concentrated among the middle and upper tail of the wage distribution, it would be evidence of a positive a selection process.

[21]In the appendix, tables 7-10 show the figures for the rural to urban/metropolitan migration pattern. The results does not affect the conclusions here presented.

Table 10: Rural Stayers structure and average education levels for individuals and their households, 2005

	Rural Stayers		
Wage group	Share (%)	Years of education	Years of household education
Low	52.48	4.97	4.95
Middle	29.24	6.17	6.11
High	18.26	7.02	7.02

All average of education expresed in years
Source:own elaboration, based on EPH 2005

First, looking at the structure concerning wage levels, migrants are principally in the middle of the wage distribution in 1997/98 and in the middle and upper tail of the distribution in 2005[22]. It is worth noting ex-post and on average, that migrants and their families are more educated than their stayer counterparts. The difference between the figures on education is considerable, particularly when looking at the average family education. The problem with this descriptive analysis is that it is highly possible that other covariables co-determine these figures and at the same time, reinforce the importance to control for the level of human capital accumulation in the household. If no control for this variable is made, then it is highly possible that the econometric specifications will be misspecified.

This descriptive information concerning education levels for individuals contrast with the finding by Richard (1996), in which migrants from Paraguay to Argentina are negative selected in relation to education.

Table 11: Poverty among movers to metro and urban/metro areas and the whole labor market in 2005

Poverty status	metro pct.	urban/metro pct.	labor market pct.
Extrem poor	3.77	3.30	11.06
Poor	5.48	15.98	17.66
Non poor	90.75	84.02	71.28

Source: own elaboration, based on EPH 2005

Table 9 offers descriptive figures concerning poverty distribution in 2005 in relation to the rural-metro migrant status. The results indicate ex-post a positive selection concerning poverty. This is because the poverty among movers is smaller than the poverty throughout the whole labour market.

[22]This table offers similar evidence founded by Chisquiar and Hanson (2005) for the Mexican migrant to the US by using counterfactual predictions of skill premia for US immigrants. However, these figures only say that migrants were highly successful ex-post in their migration adventure. These figures don't relate to anything about where they were in the wages distribution at the origin.

An additional element is provided by Otter and Villalobos (2009)[23] for 2005; in this paper, migration appears to be an equalization mechanism which does not tend to increase poverty. Quantile regressions are used to show that the rural metropolitan migration can explain wages for the 10^{th} and 25^{th} quantiles, indicating that a relative shortage of labour in metropolitan areas for those wage segments exist in 1997/98 and 2005.

The ex post evidence, that on average rural metropolitan migrants are relatively more educated and the premium for lower quantiles in the wage distribution together with lower poverty incidence among migrants can be consistent with the idea that rural metropolitan migration works as a source of additional incomes for workers with low opportunity cost, as a consequence of high levels of household education. This hypothesis will be tested in the structural probability model.

3.3 Inequality of earnings

Concerning inequality, Table 10 reports the calculated Gini coefficients for the hourly wage distributions for the studied years in rural, metropolitan and metropolitan and urban areas together.

Table 12: Gini coefficients for hourly wages in Paraguay, by areas

Year	Metro	Metro/Urban	Rural	Whole Paraguay
1997	0.447	0.470	0.472	0.493
2005	0.458	0.465	0.528	0.489

Source: own elaboration, based on EIH 1997/98 and EPH 2005

The inequality of earnings appears to be higher in rural areas than in the metropolitan ones. At country level, inequality remains at the same level of around 0.49 between 1997 and 2005. Nevertheless, rural and metropolitan areas became more unequal while the urban ones became more equal in the same time interval. If the earnings in rural and metropolitan areas are correlated enough, the standard Roy model predicts for the Paraguayan migrants from rural areas to be negative selected as unobservable in both years to metropolitan areas and negative selected to metropolitan and urban areas together only for 2005. The standard Roy model also shows that the self-selectivity of migrants is a "relative concept" because this also depends on the selection of origin and destination areas. For this reason, it is important to look at different aggregation levels for the destination and because the metropolitan area is more equal than the metro/urban area, it is also expected that the magnitude of this self-selection concerning unobservable should be higher towards the metropolitan area than the metro/urban ones.

[23] Otter and Villalobos (2009) estimated wage quantile regression controlling for all possible factors that the household survey allows. They use the quantreg package in R.

4 Theoretical framework

Migration is the result of a cost-benefit analysis, where potential migrants evaluate their comparative advantages in order to stay or move. The standard Roy model states that, given the distribution of incomes at origin and destination, migrants with higher abilities tend to migrate to more unequal distributions (areas) and vice versa. The only assumption is that the determinants of the incomes of potential migrants in home and host areas must be correlated.

The standard Roy model does not consider any switching cost and therefore important information is not taken into account if moving costs decrease with the amounts of human capital. It is reasonable to assume, that the same characteristics which yield individuals to obtain higher wages can be related to a reduction of moving cost. If this is the case, it is plausible that individuals on the top of the income distribution at home (origin) decide to move to a host area with a more equal wage distribution.

4.1 Extended Roy model

This model, presented by Borjas (1988), considers random mobility costs. The log wages at home area are described by:

$$lnW_1 = u_1 + e_1 \tag{1}$$

where u_1 is the average log wage at the home area and e_1 is the zero mean disturbance with variance σ_1^2. In the same way, define the log wages at host area (area 2) are defined such that:

$$lnW_2 = u_2 + e_2 \tag{2}$$

Both wage distributions have a joint normal distribution, where e_1 and e_2 can be interpreted as unobservable abilities of individuals.

Assuming that C represents the migration costs which, in this extension of the standard Roy model, are not fixed but a proportion of the monetary and non-monetary cost of migration as a proportion of home income. Migration occurs if $\frac{W_2 - W_1}{W_1} > C$, which is approximately $lnW_2 - lnW_1 > C$.

$$C = \gamma + \varepsilon \tag{3}$$

The extended version of the standard Roy model assumes that C is normally distributed with mean γ and error $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. With this information, an individual moves if the index function $I^* = (u_2 - u_1 - \gamma + e_2 - e_1 - \varepsilon) > 0$ and stays if $I \leq 0$. Assuming the normality conditions and defining $\sigma^v = \sqrt{Var(e_2 - e_1 - \varepsilon)}$, $z = -\frac{u_2 - u_1 - \gamma}{\sigma^v}$ and $\eta = \frac{e_2 - e_1 - \varepsilon}{\sigma^v}$, the probability to move is given by:

$$Pr(\eta > z) = 1 - \Phi(z) \tag{4}$$

where $\Phi()$ is the cumulative distribution of the standard normal and following Heckman (1979), the unobserved wage of a mover in the origin region is given by

$$E(lnW_1 | I^* > 0) = u_1 + \frac{\sigma_1 \sigma_2}{\sigma^v} \left[(\rho_{1,2} - \frac{\sigma_1}{\sigma_2}) - \rho_{1,\varepsilon} \frac{\sigma_\varepsilon}{\sigma_2} \right] \lambda \tag{5}$$

and the observed wages at destination as

$$E(lnW_2 | I^* > 0) = u_2 + \frac{\sigma_1 \sigma_2}{\sigma^v} \left[(\frac{\sigma_2}{\sigma_1} - \rho_{1,2}) - \rho_{2,\varepsilon} \frac{\sigma_\varepsilon}{\sigma_1} \right] \lambda \tag{6}$$

where $\rho_{1,2}$ represent the correlation coefficient between the disturbances e_1 and e_2 . $\rho_{1,\varepsilon}$ and $\rho_{2,\varepsilon}$ are the correlation coefficients between e_1 and the error of the moving cost ε and beet wen e_2 and ε respectively. $\lambda(z) = \frac{\phi(z)}{1 - \Phi(z)}$ is the inverse of Mill´s ratio where $\phi()$ is the standard normal density function.

Because $\lambda(z)$ and $\frac{\sigma_1 \sigma_2}{\sigma^v}$ are strictly positive, on average and given their observed characteristics, a migrant is better off than an average person in the home region if
$\left[(\rho_{1,2} - \frac{\sigma_1}{\sigma 2}) - \rho_{1,\varepsilon} \frac{\sigma_\varepsilon}{\sigma_2} \right] > 0$ and better off than an average person in the host region if
$\left[(\frac{\sigma_1}{\sigma 2} - \rho_{1,2}) - \rho_{2,\varepsilon} \frac{\sigma_\varepsilon}{\sigma_1} \right] > 0$. By using simple algebra, the above conditions can be expressed respectively as

$$\frac{\sigma_1}{\sigma_2} > \frac{1}{\rho_{1,2}} + \frac{\rho_{1,\varepsilon}}{\rho_{1,2}} \frac{\sigma_\varepsilon}{\sigma_1} \tag{7}$$

$$\frac{\sigma_1}{\sigma_2} > \rho_{1,2} + \rho_{2,\varepsilon} \frac{\sigma_\varepsilon}{\sigma_1} \tag{8}$$

The second term in equation (7) helps to determine the direction of selection bias of the migrants in relation to an average person in the home region. Given variance of earnings in the home and host regions and the observed characteristics, the higher the correlation between unobservable of

wages in home region and unobservable determinants of moving cost, the higher the probability that migrants could be selected from the upper tail of the home earnings distribution.

Brücker and Trübswetter (2007) decompose the second terms in (5) and (6) in order to identify composition and scale effects[24] (Borjas, 1988) and formulates the following prediction of this model:

- Increasing the inequality of earnings at home as well in the host region has an ambiguous effect on the selection bias of the migrant population
- Increasing the difference in earnings between the host and home regions reduces the selection bias of the migrant population
- Increasing the moving cost increases the selection bias of the migrant population
- The higher the correlation between unobserved labour market abilities and moving cost, the higher the possibility that the migrants are positive selected
- The higher the correlation between earning in home and host regions, the higher the selection bias of the migrant population
- The higher the negative correlation between earnings and moving costs, the smaller the selection bias if the net difference in earnings among regions is positive

The extended version of the Roy model reveals important implications about the selection biases in terms of unobserved and observed characteristics. Again, Borjas (1988) shows that positive selection in observed characteristics (i.e schooling) occurs if the labour market in the host region attaches a higher return than in comparison to the home region. He addresses the idea that the selection condition of observed characteristics has nothing to do with the selection in unobserved characteristics. In other words, the two types of selection are independent from each other. The estimation strategy presented in the next section controls for observed characteristics as well for the impact of unobservable on wages in order to estimate a structural probability model for rural metropolitan migration.

[24]The composition effect measures how a change in the ability mix of a constant-sized immigrant pool affects the selection bias, holding the size of the flow constant. The scale effect captures what happens to selection bias as the size of the flow is increased for any given mix of abilities.

4.2 Estimation

The extended Roy model presented in this paper finds a suitable counterpart in a switching regression model, presented by Goldfeld and Quandt (1973) with endogenous switching[25] (Maddala and Nelson, 1975; Maddala, 1983).

In the first step, a reduced form for the probability to move was estimated, and by doing so, the inverse Mills ratio can be obtained and used in a Maximum Likelihood wage regression. This information is later used to predict wages at the destination for rural stayers and wages in rural areas for movers with selectivity control, and in this way the potential wage differential is calculated, which is later used in a structural probability model.

Equation (1) and (2) can be rewritten as

$$lnW_{1i} = X_{1i}\beta_1 + \mu_{1i} \tag{9}$$

$$lnW_{2i} = X_{2i}\beta_2 + \mu_{2i} \tag{10}$$

where X_i is a vector of personal characteristics determining wages. In the same way, the index function can be represented for the *ith* individual as

$$I_i = \delta(lnW_{2i} - lnW_{1i}) - Z_i\psi - \varepsilon_i \tag{11}$$

where additionally the migration cost $Z_i\psi + \varepsilon_i = C$ (counterpart of (3)) depends again on personal characteristics in Z_i and unobservable ε_i. In order to identify this system, at least one variable in Z_i must be not included in X_i. The index function cannot be estimated in a structural form because $lnW_{2i} - lnW_{1i}$ is endogenous.

To solve this endogeneity problem, Lee (1978) and Willis and Rosen (1979) propose a three step strategy. A reduced form of the index function can be estimated by using a probit Maximum Likelihood estimator where $I_i = 1$ if $I^*> 0$ and $I_i = 0$ otherwise.

$$I_i^* = \delta(X_{2i}\beta_2 - X_{1i}\beta_1) - Z_i\psi + \delta(\mu_{2i} - \mu_{1i}) - \varepsilon_i = Z_i^*\psi^* + \varepsilon_i^* \tag{12}$$

[25] Using the same notation, endogenous switching satisfies that $(\mu_{1i}, \mu_{2i}, \varepsilon_i) \sim N(0, \Sigma)$, where $\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{1\varepsilon} \\ \sigma_{12} & \sigma_{22} & \sigma_{2\varepsilon} \\ \sigma_{1\varepsilon} & \sigma_{2\varepsilon} & 1 \end{bmatrix}$ is the Covariance Matrix.

The parameter vector $\hat{\psi}^*$ can be suitably estimated and therefore the inverse Mills' ratio for stayers and migrants can be calculated. Equations (9) and (10) can be estimated for each individual and consequently a potential differential wage can also be calculated. Both equations can be estimated for stayers and movers with control for selection bias. For stayers

$$lnW_i = X_i\beta_1 - \sigma_{e1\eta}\frac{\phi(Z_i^*\hat{\psi}^*)}{\Phi(Z_i^*\hat{\psi}^*)} + \mu_{1i} \tag{13}$$

and for movers

$$lnW_i = X_i\beta_2 - \sigma_{e2\eta}\frac{\phi(Z_i^*\hat{\psi}^*)}{1-\Phi(Z_i^*\hat{\psi}^*)} + \mu_{2i} \tag{14}$$

This paper uses the Heckman Maximum Likelihood estimator for survey data which also takes into account the correlation between primary sample units avoiding the underestimation of standard errors and consequently avoiding the overestimation of sample selection bias or possible self-selectivity of migrants[26].

The last step to estimate structurally the index function is by using the estimated potential wage differential and other variables as regressors which determine migration probability but not through wages. Unemployment rates, average family education and the number of children in household were used, as were the excluded variables in (13) and (14) to identify (12).

[26]The maximum likelihood method has been shown to produce consistent estimates under a few plausible conditions. Maximum likelihood estimates have the further advantage of being normal and efficient if sample sizes are large enough (Gujarati, 2003: 113).

5 Estimation Results

First, the strategy seeks to obtain consistent estimates of the individual probability to migrate, by using these estimates to control for the residual correlation between the selection equation and the wages equation. Unbiased estimates for wages in Rural, Metropolitan and Urban/Metropolitan areas were obtained. Lastly, the estimated potential wage differential, which is not more endogenous, is included as a regressor in the structural probit for migration. District aggregated lagged values of the explanatory variables of interest which were obtained from an available past survey with similar characteristics were used, due to the fact that the decision to migrate was made before the studied cohort year. Lagged values for 2005 are obtained from *EIH 2000/2001* and for 1997/98 from *EPH 1992*.

5.1 Probit, reduced form

Following Harris and Todaro (1970)[27] beyond the potential differential wage, personal characteristics such as education, potential experience, gender and ethnic origin jointly with household composition variables such as marital status and children in household are expected to influence the probability to move. Unemployment rates could impact the probability to be employed and in this way affect expected earnings.

Tables 9 and 10 in the appendix show the probit reduced forms for the probability to migrate to metropolitan and to urban/metropolitan areas in 1997 and 2005. The probit regressions indicate a negative relationship between educational progress and the probability to move. This effect is stronger for individuals with college education but can also reflect that those people are not willing to migrate because they are permanently employed in profitable farming firms or in the public sector. This finding is contradictory with the argument that more educated people are in a better position to get information about labour markets and take rational decisions concerning migration (Linnaman and Graves, 1983). However, Paraguay appears to be a country where contradictory findings find rational explanations. Richards (1996) and Molinas (1999) find a negative tendency for more educated individuals to move. Richard argues that workers with relatively high educational levels have comparatively good prospects to get a job in their areas of origin[28].

[27] Todaro assumes that the person estimates the net present value of the migration decision in a continuous scale: $V_u(0) = \sum_{i=0}^{n} p(t)Y_u(t)e^{-\rho t}dt - c(0)$ where $Y_u(t)$ is the real income in the urban sector at time t, $p(t)$is the probability to get a job in the urban sector at time t, $C(0)$is the migration cost.

[28] Richard (1996) studies the Urban-Urban Migration pattern between Asunción, Paraguay and Buenos Aires, Argentina. This result could be valid for the rural metropolitan migration pattern inside Paraguay due to the fact that the lack of opportunities between Asunción and Buenos Aires can be compared to some degree with the lack of opportunities between rural areas and Asunción. Molinas (1999) estimates probability models without decomposition in

This paper finds that marriage is associated with a reduction in the probability to migrate, but this effect appears to be significant only in 1997. The number of children per household reduces the probability to move as factor cost, again this result is consistent for Paraguay with the results of Richards (1996) and Molinas (1999). As expected, the average household education affects positively the probability to move by reducing the probability that in the case of migration a household will suffer a considerable reduction in welfare. Consequently, it is in accordance with the idea that the cost in the case of migration will be somehow proportional to the household welfare dependence on any household member. For this reason the higher the average household education, the lower the dependence on an individual household member and therefore, the higher the probability to move for the potential rural metropolitan migrant..

An ethnic variable called *"Castellano"* has a strong and significant pro migration impact. This is a design variable equal to the unity when the individual exclusively speaks *Castellano* (Spanish) and is therefore supposed to have a non-aboriginal origin. As expected, under the consideration that imperfect labour market information takes place in rural areas, the Lagged departmental unemployment rates at origin appear to stimulate migration. Potential experience appears to have a significant and negative non-linear impact on migration probability[29].

The mentioned output controls for interior and border departments[30] and as expected the geographic location within Paraguay plays an important role when people choose a migration destination, especially given the differential of amenities between Paraguay, northern Argentina and southern Brazil[31]. WWorkers who originate from border departments are less likely to move towards Asunción. This finding is also expected because Asunción represents a relatively expensive alternative in comparison to moving to Misiones and Corrientes in Argentina given the higher development gap between the rural areas and the aforementioned destinations[32]. In general, border regions have been characterized by weak domestic markets and high transaction costs[33].

migration patterns, but again, the high employability of relatively high skilled workers in rural productive units called *"estancias"* could not only explain this finding but also explain the more unequal wage distribution in rural areas.

[29]This finding is also present in Richards (1996) and Molinas (1999) in terms of age. Nevertheless, potential experience is highly correlated with age.

[30]Border departments: *Amambay, Canindeyú, Alto Paraná, Itapúa.* Interior departments: *San Pedro, Caaguazú, Paraguarí.*

[31]Fafchamps and Shilpi (2008) have shown that migrants are concerned with their welfare relative to that of their home district as well as to that in their potential destination.

[32]The human development index in 1995 for the Argentine province of Misiones is 0.840, and 0.971 for Corrientes (P.A.DH. Informe Argentino sobre Desarrollo Humano 1998; Bs. As.), while the respective values in 2006 for the Paraguayan departments were Amambay 0.706, Canindeyú 0.742, Alto Paraná 0.744 and Itapúa 0.748. For instance, Asunción had an index of 0.837 (Atlas de Desarrollo Humano Paraguay 2007 PNUD Paraguay). From these development indicators, it can be estimated that the development gap between Asunción and Misiones is around 11 years and the gap between the Paraguayan border departments and Misiones is around 11 years plus 0.1 HDI points.

[33]See, Henk van Houtum, Uncertainty and dependency in cross border business movements, Regional Science Asso-

Nevertheless, border regions in Paraguay still attract migrants because of the capture of exogenous economic dynamics which continues driving the Paraguayan growth process (Masi et al, 2000). In general, almost all results hold when the reduced probability models for the rural to urban/metropolitan migration are estimated.

5.2 Heckit wage regressions

Tables 11, 12, 13 and 14 in the appendix show the wage regressions based on a Mincer structure for rural stayers and rural movers to metropolitan and to urban/metropolitan areas in 1997 and 2005. All regressions control for the correlation between unobservable characteristics from the migration index function and the wage equation. The probability model corresponds exactly to the probit reduced form presented above. The excluded variables in the wage regression are the number of children per household, marital status and the average family education in years. These variables should not be correlated with the residual of the wage regressions, which is a highly plausible assumption. These exclusion restrictions are made for identification purposes[34].

The log hourly wages are explained by human capital accumulation indicators such as education grades and potential experience. These Mincer variables have the expected magnitudes, significance and directions. Economic environment variables such as the log of the lagged district mean household income at origin, informal sector, occupation, and economy sector show the expected impact on wages. The regressions also control for public sector and department where the individual is employed.

Based on the probit reduced form specification, it is possible to examine whether the second terms in equations (13) and (14) are significant positively or negatively different from zero. The sign of the Lambdas indicates whether the migrant/stayers are positively or negatively selected in relation to unobserved characteristics. Regarding unobserved characteristics, strong evidence has been found that in 1997 rural metropolitan movers were negative selected while with urban/metropolitan areas this evidence is weaker. For 2005 there is no evidence about the self-selectivity of movers. In unobservable, rural metropolitan migrants appear to be a random sample

ciation, Napa, USA.

[34]To adequately and accurately correct for the impact of sample selection, some measure is required to control for the colinearity between the regressors in the wage equation and the correction term. The most effective way of doing this is by incorporating at least one variable into the probabilistic equation that is not contained in the wage equation. This variable needs to influence the individual's likelihood of being employed, but have little or no impact on wages. Following Puhani (2000: 58), household variables are likely to be most apt for use as exclusion restrictions in labour market analysis, since these variables are more likely to affect the employability prospect without also affecting the wage an individual would attain. This is not the case for most other variables, especially those that denote personal characteristics, since these are usually correlated with the wage function.

of the rural labour force.

The negative selection in unobservable holds for 1997 at the 15% level; this indicates that the results are not dramatically driven by the geographical nature of the destination and apply also for urban destinations across Paraguay.

5.3 Probit, structural form

By estimating the wage regressions with sample bias controls for stayers and movers, the potential wage differential can be calculated. The structural forms, in Tables 15 and 16 also include as regressors marital status, average household education, lagged district unemployment rates, and controls for border and interior departments.

The results are highly robust and almost all variables for both years have expected signs and statistical significance. The potential wage differential appears to be an important determinant of the probability to move. The wage differential appears to have a higher impact when migrants move to a metropolitan area than when moving to urban/metropolitan areas. This result confirms the selectivity issue involved in the migration process.

Marital status appears to recover its theoretical importance determining probability of migration. As expected, the higher the lagged departmental unemployment rates at origin, the higher the propensity to move. This effect is reported to be higher in 1997 than for 2005.

The effect of the controls for border and interior departments tends to disappear when the destination regions are comprised of metropolitan and urban areas in comparison with only metropolitan area as a destination region. This result is expected considering the concentration of the metropolitan area (Asunción) relative to the distribution of urban areas across Paraguay. This result also applies for the reduced form estimations.

Average household education appears to encourage migration according to the idea that the cost incurred during migration will be proportional to the household welfare dependence of any household member. This result is not surprising and reveals the necessity to consider that in all migration process, the household education level matters.

Following Richards (1996), it is not always clear when discussing the decision to migrate exactly what constitutes the decision-making unit. Is the decision to move to an alien environment (a

necessarily traumatic one) a purely individual choice, or is it a collective and, hence, more complicated process? The calculation of the net marginal benefits of migration are clearly more complicated if the decision-making unit is a married couple. Not only is the cost of migration increased substantially but (especially if both partners work) the expected return is more uncertain.

The rationale behind the migration decisions is that, potential migrants are less able to migrate if they are part of more complex household structure. Singles are more able to migrate than couples. The same situation occurs regarding the number of children in household. The household composition reflects the fact that each of these represents an opportunity cost (monetary or spiritual) and gives each household member a certain degree of freedom.

Household composition matters not only for the composition itself, but also for their intrinsic capacity to generate welfare. Households with lower welfare prospects should give their members a lesser degree of freedom or higher opportunity cost than households with higher levels of welfare. The argument mentioned above is very important when empirical studies about the determinants of migration are made.

First, it is very important to control for household composition and welfare prospects. A single measurement such as household size can be inappropriate because it is usually highly correlated with the number of children within the household. Nevertheless, the number of children is a better way to control for "less degree of freedom" because it is understood that children reduce the welfare prospects by reducing the average years of education or because they are unable to obtain minimum levels of income to significantly increase the household's welfare. Second, controlling for the household's mono-parental condition is desirable because of the fact that, in mono-parental households, household heads have typically more difficulties to move when they have children to care for. Third, these difficulties can be somehow overcome when there is a high intra-household welfare generation capacity. This welfare generation capacity was successfully instrumented by the average household years of education, which is highly robust for all reduced and structural specifications.

Given that rural metropolitan migrants are, on average, relatively more educated than their stayer counterparts, the premium for rural metropolitan migration for lower quantiles in the wage distribution[35] and the lower ex-post poverty incidence among rural metropolitan migrants, it is not inconceivable to think that this picture can be consistent with the idea that rural metropolitan migration works as a mechanism to obtain additional incomes for low opportunity cost workers, as a consequence of relatively high levels of household education.

[35] Otter and Villalobos (2009)

5.4 Unemployment and probability models

In this analysis, unemployed was excluded from the regressions, which may bias the estimates if unemployment rates in the group of stayers and movers as their characteristics are different. As the unemployment rates and their characteristics slightly differ between both groups, it is possible that my results are slightly biased regarding the whole labor market. However, including the unemployed in the probit regressions does not change the results much: the parameters have a similar size and the same pattern of significance is observed[36]. For this reason, I can conclude that the exclusion of the unemployed has not affected the main findings of this paper.

[36]The results are available from the author upon request. When probability models are estimated with unemployment, only one important change is observed. The coefficient for the potential wage differential in 1997/98 goes from 0.629 to 0.954 at 1% level.

6 Conclusions

Since the inequality of wages in rural areas was higher than in metropolitan areas for 1997 and 2005, the standard Roy model predicts that the migrants were negatively selected. An extended Roy model was used to incorporate the possibility that moving costs could be correlated with skills in labour markets. To empirically investigate this issue, a switching regression model with endogenous switching was implemented.

Concerning unobservable, the results show strong evidence that migrants were negatively selected in 1997 to metropolitan areas. For 2005, no evidence was found regarding the self-selectivity of migrants. The findings of this study show that rural metropolitan and rural urban/metropolitan migrations do not represent a loss of the relative human capital stock for rural areas in relation with the metropolitan or urban/metropolitan ones (because of the non-positive migrant selection). For 2005, rural movers to metropolitan or urban/metropolitan areas appear to be a random sample of the Paraguayan rural labour force. The standard Roy model appears to fail in properly predicting the selection bias of unobserved characteristics. Regarding the observed characteristics, rural metropolitan migrants seem to be negative selected in both 1997 and 2005, but when the destination to urban/metropolitan areas was expanded, the selection concerning basic and secondary education tends to disappear in both years.

Concerning the determinants of the rural metropolitan and rural urban/metropolitan migration, a structural probit model was estimated including the estimated potential wage differential for each individual, marital status, number of children per household and the average household education. Evidence has been found supporting that the potential wage differential determines the probability to move. The higher the wage differential, the higher the probability to move. Marital status and children per household also affect the migration probability in the expected ways as factor costs.

Household composition matters not only for the composition itself, but also for their intrinsic capacity to generate welfare. Households with lower welfare prospects should give their members a lesser degree of freedom or higher opportunity cost than households with higher levels of welfare. The argument mentioned above is very important when empirical studies about the determinants of migration are made. In this study, this argument is endorsed by the fact that the average household education is an important and a structural determinant of the probability of migration. The degree to which the potential remaining members of the household can generate income is also important information for the potential migrant. Households with lower levels of human capital accumulation are more dependent of the potential monetary contribution of each

member to the household. The cost in the case of migration will be somewhat proportional to the household welfare dependence of any household member and is therefore not surprising that higher levels of average education imply a higher degree of freedom for the individuals by reducing the costs in case of migration.

This paper provides important evidence in order to reveal the dual impact of education on rural metropolitan migration probability. On the one hand, individuals own education discourages migration; on the other, the average household education encourages migration by increasing the degree of freedom of the potential migrant. Ignoring this dual impact of education can yield to obtain biased impacts of education on migration and therefore the effect of unobservable on migration probability could be misleading.

The fact that rural metropolitan migrants are, on average, relatively more educated than their stayer counterparts, the founded premium for rural metropolitan migration for lower quantiles and the lower ex-post poverty incidence among rural metropolitan migrants, can be consistent with the idea that rural metropolitan migration works as a mechanism to obtain additional incomes for low opportunity cost workers, due to the fact that for the most part they come from relatively higher educated families.

Rural metropolitan migration seems to be an equalization mechanism with desirable consequences in terms of employment and wages. It can be argued that reducing migration costs to metropolitan areas appears to be consistent with the idea of achieving more flexible labour markets, contributing to reduced unemployment, increasing lowest wages and therefore improve the income distribution by facilitating a low opportunity cost migrant to obtain a significant source of income.

References

Borjas, G. (1987): "Self-Selection and the Earnings of Immigrants", *American Economic Review*, American Economic Association, vol. 77(4), pages 531-53, September.

Borjas, G. (1988): "Immigration And Self-Selection", *NBER Working Papers* 2566, National Bureau of Economic Research, Inc.

Brücker, H. and Trübswetter, P. (2007): "Do the Best Go West? An Analysis of the Self-Selection of Employed East-West Migrants in Germany", *Empirica*, Springer, vol. 34(4), pages 371-395, September.

CEPAL. (1998): "Impacto de la crisis asiática en América Latina". Santiago de Chile.

Fields, G. (1975): "Rural-Urban Migration, Urban Unemployment, and Underemployment, and Job Search Activity in LDC's", *Journal of Development Economcis*, 2, 2, 165-38.

Galeano, L. (1997): "La Migración Interna en el Paraguay", Dirección General de Estadísticas, Encuestas y Censos, Paraguay.

Goldfeld, S. and Quandt, R. (1973): "The Estimation of Structural Shifts by Switching Regressions", *Annals of Economic and Social Measurement*, 2, 475-485.

Greenwood, M. J. (1975): "Research on Internal Migration in The United States: A Survey", *Journal of Economic Literature*, 13, 397-433.

Greenwood, M. J. (1985): "Human Migration: Theory, Models, and Empirical Studies", *Journal of Regional Science*, 25(4), 521-544.

Greenwood, M. J. (1995): "Internal Migration in Developed Countries". *Handbook of Families and Population Economics*. Rosenzweig and Stark (Eds.). North-Holland, Amsterdam.

Gujarati, D. (2003). Basic Econometrics. 4th ed. Boston: McGraw-Hill.

Harris, J. and M. Todaro (1970): "Migration, Unemployment, and Development: a Two-Sector Analysis", *American Economic Review*, 60, 126-42.

Heckman, J. (1979): "Sample Selection Bias as a Specification Error", *Econometrica,* 47, pp.153-161.

Hosmer, D. W., and S. Lemeshow (2000): Applied logistic regression. John Wiley and Sons, New York, New York, USA.

Lewis W.A. (1954): "Economic development with unlimited supplies of labour", *The Manchester School of Economic and Social Studies*, vol. XXII, N.2, May.

Linneman, P. and P. Graves(1983): "Migration and Job Change: a Multinomial Logit Approach", *Journal of Urban Economics,* 14, 263-79.

Maddala G. (1983): Limited-dependent and qualitative variables in econometrics. Cambridge, England: Cambridge University Press.

Masi, F., R. Penner, R. Dietze. (2000): "Evaluación del Rol de las Regiones Fronterizas en el Proceso de Desarrollo Económico del Paraguay", Departamento de Economía Internacional Banco Central del Paraguay.

Mincer, J. (1994): "Human Capital: a Review", *Labor Economics and Industrial Relations*, edited by C. Kerr and P. Staudohar, Harvard University Press.

Molho, I. (1986): "Theories of Migration: A Review", *Scottish Journal of Political Economy*, 33(4), 396-419.

Molinas, J. (1999): "Migración Interna en Paraguay: Un Análisis Económico de la Encuesta de Hogares 1996", Mejoramiento de las Encuestas de Condiciones de Vida en America Latina y el Caribe, Programa MECOVI-Paraguay.

Otter, T. and C.Villalobos, (2009) "Rural-Urban Migration and Wages in Paraguay", Draft. Georg August Universität Göttingen.

Parrado, E. and M. Cerrutti (2003) "Labor Migration between Developing Countries: The Case of Paraguay and Argentina", *International Migration Review*, Vol. 37 Number 1 (spring 2003):101-132.

PNUD. (1998) : "Informe Argentino sobre Desarrollo Humano" Buenos Aires, Argentina.

PNUD. (2007): "Atlas de Desarrollo Humano Paraguay", Asunción, Paraguay.

Puhani, P. (2000): "The Heckman Correction for Sample Selection and its Critique", J*ournal of Economic Surveys.* 14, 1: 53-68.

Richards, D. (1996), "Household Migration in the Southern Cone: The Case of Paraguay", *Applied Economics* 28: 89-95.

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

Schultz, T. (1961): "Investment in Human Capital", *American Economic review,* 51, 1-17.

Sosa Z. (1996): "La Migración Femenina en Paraguay", Dirección General de Estadísticas, Encuestas y Censos, Paraguay.

Tannuri, M.E. (2000): "Three Essays on the Economics of Immigration to the US", PhD Dissertation , University of Illinois at Urban Champaign.

Tannuri-Pianto, M. and D. Pianto (2002): " Informal employment in Brazil – a choice at the top and segmentation at the bottom: a quantile regression approach", Working paper n 236, Department of Economics, University of Brasilia.

Todaro, M.P. (1969): "A Model of Labor Migration and Urban Unemployment in LDCs". *American Economic Review* 59: 138-148.

Van Dijk, J., H. Folmer, H.W. Herzog Jr. y A.M. Schlottmann (1989b): "Labor Market Institutions and the Efficiency of Inter regional Migration: A Cross-Nation Comparison", *Migration and Labor Market Adjustment*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Yashiv, E. (2004): "The Self-Selection of Migrant Workers Revisted": *IZA Discussion Papers 1094*, Institute for the Study of Labor (IZA).

Table 1: Characteristics of Rural Origin Heads of Households, 1997/98.

	Rural-urban Mig.		Rural-met Mig.		Rural Non-Mig.	
Variables	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Family Monthly Income	1,035,007	910,894	1,028,530	1,052,743	749,594	835,937
Female ratio	0.20	0.40	0.21	0.41	0.17	0.38
Age	42.80	17.06	32.58	15.14	48.36	15.81
Household Size	4.98	2.61	2.69	1.94	5.16	2.62
Married (%)	0.77	0.42	0.51	0.50	0.81	0.39
Years of Education	5.56	3.81	6.40	3.28	4.08	2.90
Income (PPP)	3,290	2,990	5,179	2,431	2,241	2,891
Infant Mortality Rate	40.63	12.84	32.56	7.96	39.76	13.51
Economic Act. Pop.	0.78	0.42	0.92	0.28	0.86	0.35
Unemployed	0.05	0.21	0.02	0.14	0.14	0.11
Formal	0.33	0.47	0.40	0.49	0.14	0.34
Informal	0.67	0.47	0.60	0.49	0.86	0.34
Observations	65		34		1,751	

Source: own elaboration, based on EIH 1997/98

Table 2: Characteristics of Urban/Metropolitan non-migrants Heads of Household, 1997/98

	Urban non-migrant		Metro non-migrant	
Variables	Mean	Std. Dev.	Mean	Std. Dev.
Family Monthly Income	1,581,170	2,399,737	2,623,416	2,759,868
Female ratio	0.23	0.43	0.29	0.45
Age	46.08	15.45	46.54	14.87
Household Size	4.61	2.24	4.54	2.29
Married (%)	0.79	0.41	0.72	0.45
Years of Education	6.21	3.98	8.36	4.58
Income (PPP)	4,605	6,278	8,026	9,112
Infant Mortality Rate	43.97	13.78	32.58	8.65
Economic Act. Pop.	0.81	0.39	0.82	0.39
Unemployed	0.03	0.17	0.02	0.15
Formal	0.25	0.43	0.35	0.48
Informal	0.75	0.43	0.65	0.48
Observations	939		896	

Source: own elaboration, based on EIH 1997/98

Table 3: Characteristics of Rural Origin Heads of Households, 2005.

	Rural Mig. to Urban		Rural Mig. to Met.		Rural Non-Mig.	
Variables	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Family Monthly Income	1940618	1977737	1496868	770353	1955731	2089841
Female ratio	0,35	0,48	0,26	0,44	0,22	0,41
Age	43,92	17,62	45,84	15,16	49,13	15,69
Household Size	4,26	2,63	2,26	1,53	4,76	2,46
Married (%)	0,80	0,40	0,68	0,47	0,76	0,43
Years of Education	7,68	5,57	8,15	3,50	5,33	3,95
Income (PPP) (4642)	4279	4955	7034	5890	3833	4488
Infant Mortality Rate	38,08	8,91	28,36	6,19	31,58	11,85
Economic Act. Pop.	0,81	0,40	0,75	0,43	0,87	0,33
Unemployed	0,04	0,19	0,01	0,10	0,01	0,12
Formal	0,43	0,50	0,47	0,50	0,21	0,59
Informal	0,57	0,50	0,53	0,50	0,79	0,41
Observations	39		19		1,566	

Source: own elaboration, based on EPH 2005

Table 4: Characteristics of Urban/Metropolitan non-migrants Heads of Household, 2005

	Urban non Migrant		Metro. non Migrant	
Variables	Mean	Std. Dev.	Mean	Std. Dev.
Family Monthly Income	2,463,150	3,161,740	3,344,164	5,891,360
Female ratio	0.29	0.45	0.35	0.48
Age	45.84	15.16	47.30	15.19
Household Size	4.32	2.27	4.28	2.12
Married (%)	0.71	0.45	0.69	0.46
Years of Education	7.13	4.71	8.55	4.95
Income (PPP)	5,271	5,890	6,737	11,412
Infant Mortality Rate	36.58	10.97	27.73	7.82
Economic Act. Pop.	0.82	0.38	0.81	0.39
Unemployed	0.03	0.16	0.03	0.17
Formal	0,41	0,51	0,51	0,50
Informal	0.59	0.49	0.49	0.50
Observations	939		896	

Source: own elaboration, based on EPH 2005

Table 5: All migration flows, by origin and destination, 1997/98 (heads only)

Destination / Origin	N° of Migrants	% of destination flow	% of origin flow	% of total flow
Migrants to Rural Areas				
Rural	23,417	49.23	51.30	16.95
Urbano	17,050	35.85	36.37	12.34
Metropolitano	7,095	14.92	15.56	5.14
Total	47,562	100.00	34.44	34.44
Migrants to Urban Areas				
Rural	11,267	29.86	24.68	8.16
Urbano	16,129	42.75	34.40	11.68
Metropolitano	10,332	27.39	22.67	7.48
Total	37,728	100.00	27.32	27.32
Migrants to Metropolitan Area				
Rural	10,961	20.75	24.01	7.94
Urbano	13,706	25.95	29.23	9.92
Metropolitano	28,157	53.30	61.77	20.39
Total	52,824	100.00	38.25	38.25
Migrants Rural, Urban and Metropolitan Areas				
Rural	45,645	33.05	100.00	33.05
Urbano	46,885	33.95	100.00	33.95
Metropolitano	45,584	33.00	100.00	33.00
Total	138,114	100.00	100.00	100.00

Source: own elaboration, based on EPH 2005

Table 6: All migration flows, by origin and destination, 2005 (heads only)

Destination / Origin	N° of Migrants	% of destination flow	% of origin flow	% of total flow
Migrants to Rural Areas				
Rural	23,928	42.40	67.29	17.54
Urbano	20,105	35.62	44.82	14.74
Metropolitano	12,403	21.98	22.14	9.09
Total	56,436	100.00	41.36	41.36
Migrants to Urban Areas				
Rural	6,066	24.72	17.06	4.45
Urbano	12,342	50.29	27.51	9.05
Metropolitano	6,135	25.00	10.95	4.50
Total	24,543	100.00	17.99	17.99
Migrants to Metropolitan Area				
Rural	5,563	10.03	15.65	4.08
Urbano	12,415	22.38	27.67	9.10
Metropolitano	37,486	67.59	66.91	27.47
Total	55,464	100.00	40.65	40.65
Migrants Rural, Urban and Metropolitan Areas				
Rural	35,557	26.06	100.00	26.06
Urbano	44,862	32.88	100.00	32.88
Metropolitano	56,024	41.06	100.00	41.06
Total	136,443	100.00	100.00	100.00

Source: own elaboration, based on EPH 2005

Table 7: Rural-Urban/Metro Migrant strucure and average education levels for individuals and their households, 1997/98

	Rural-Urban/Metro migrants		
Wage group	Share(%)	Years of education	Years of household education
Low	30.16	4.65	3.81
Middle	53.96	6.35	6.25
High	15.88	7.63	7.27

All average of education expresed in years

Source:own elaboration, based on EIH 1997/98

Table 8: Rural Stayer strucure and average education levels for individuals and their households, 1997/98

	Rural stayers		
Wage group	Share(%)	Years of education	Average household education
Low	56.10	4.22	3.37
Middle	27.67	5.49	4.19
High	16.23	7.35	5.01

All averages of education are expresed in years

Source:own elaboration, based on EIH 1997/98

Table 9: Rural-Urban/Metro Migrant strucure and average education levels for individuals and their households, 2005

	Rural-Urban/Metro migrants		
Wage group	Share (%)	Years of education	Years of household education
Low	22.59	5.75	6.67
Middle	50.08	8.39	8.87
High	27.33	9.45	9.70

All average of education expresed in years

Source:own elaboration, based on EPH 2005

Table 10: Rural Stayers structure and average education levels for individuals and their households, 2005

	Rural Stayers		
Wage group	Share (%)	Years of education	Years of household education
Low	52.48	4.97	4.95
Middle	29.24	6.17	6.11
High	18.26	7.02	7.02

All average of education expresed in years

Source:own elaboration, based on EPH 2005

Table 11: Probit reduced forms, probability to move to metropolitan area 1997-2005, Selection equation from the Heckit ML wage regressions

Variable	Coeff 97	Std.err	Coeff 05	Std.err
Secondary education	-0.611**	(0.238)	-0.338	(0.249)
College education	-1.507**	(0.587)	-1.674***	(0.508)
Married	-0.313*	(0.176)	-0.282	(0.255)
Female	0.287*	(0.159)	0.276	(0.234)
Ethnic variable - Castellano	1.008***	(0.255)	1.287***	(0.274)
Potential experience	-0.0150	(0.0279)	-0.122***	(0.0254)
Squared potential experience	-0.000364	(0.000608)	0.00201***	(0.000470)
Average family education	0.211***	(0.0425)	0.172***	(0.0544)
Children	-0.330***	(0.0710)	-0.611***	(0.0942)
Lagged Unemployment rate	0.0451***	(0.0135)	0.0172*	(0.0101)
Interior departments	0.133	(0.184)	-0.683**	(0.324)
Border departments	-0.989***	(0.337)	-1.292***	(0.426)
Constant	-3.148***	(0.738)	-1.308**	(0.600)
F	8.63		8.05	
Prob > F	0.0000		0.000	
Observations	1941		1185	
Population Size	452411		326530	
Design df	511		226	
Reference Category	Until basic education, Monoparental Household Male, Aboriginal origin, Cordillera and the rest			

Significance levels : $*$: 10% $**$: 5% $***$: 1%

Table 12: Probit reduced forms, probability to move to metropolitan and urban areas 1997-2005, Selection equation from the Heckit ML wage regressions

Variable	Coeff 97	Std.err	Coeff 05	Std.err
Secondary education	-0.144	(0.133)	-0.268	(0.170)
College education	-0.873**	(0.350)	-1.230***	(0.303)
Married	-0.274*	(0.164)	-0.152	(0.165)
Female	0.150	(0.103)	0.402***	(0.133)
Ethnic variable - Castellano	0.493**	(0.199)	0.594***	(0.185)
Potential experience	-0.0251*	(0.0148)	-0.107***	(0.0192)
Squared potential experience	0.0000516	(0.000256)	0.00181***	(0.000360)
Average family education	0.129***	(0.0281)	0.136***	(0.0341)
Children	-0.0842***	(0.0312)	-0.266***	(0.0740)
Lagged Unemployment rate	0.0337***	(0.0111)	0.00724	(0.00645)
Interior departments	0.295*	(0.169)	-0.223	(0.241)
Border departments	-0.0552	(0.203)	-0.101	(0.192)
Constant	-2.366***	(0.543)	-1.088***	(0.391)
F	7.66		7.87	
Prob > F	0.0000		0.000	
Observations	1981		1235	
Population Size	464572		338432	
Design df	567		258	
Reference Category	Until basic education, Monoparental Household Male, Aboriginal origin, Cordillera and the rest			

Significance levels : $*$: 10% $**$: 5% $***$: 1%

Table 13: Heckit Wage Reg. for rural stayers 1997-2005 (with control for migration to metropolitan area)

Variable	Coeff 97	Std.err	Coeff 05	Std.err
Secondary education	0.243***	(0.0737)	0.223***	(0.0813)
College education	0.429***	(0.149)	0.308**	(0.142)
Log of district household mean income at origin	0.339***	(0.0713)	0.610***	(0.0605)
Informal sector	-0.463***	(0.0615)	-0.268***	(0.0674)
Potential experience	0.0199***	(0.00657)	0.0231***	(0.00699)
Squared potential experience	-0.000405***	(0.000113)	-0.000381***	(0.000134)
Female	-0.412***	(0.0498)	-0.490***	(0.0602)
CEOs, Professionals, Technicians	0.740***	(0.131)	0.457**	(0.187)
Clercks, Operators, Office workers	0.455***	(0.0606)	0.328***	(0.0632)
Ethnic variable - Castellano	0.0983	(0.141)	0.171	(0.112)
Public sector	0.300**	(0.131)	0.309	(0.188)
Electricity sector	0.0116	(0.336)	0.846**	(0.390)
Finance, commerce and Transport sector	0.174**	(0.0691)	0.121*	(0.0687)
San Pedro	-0.0332	(0.156)	-0.0892	(0.166)
Cordillera	0.548***	(0.141)	-0.218	(0.161)
Guairá	0.380**	(0.185)	-0.436*	(0.239)
Caaguazú	0.0452	(0.151)	-0.189	(0.154)
Caazapá	0.381**	(0.170)	-0.458*	(0.277)
Itapúa	0.605***	(0.141)	-0.346**	(0.162)
Misiones	0.0619	(0.224)	-0.450*	(0.236)
Paraguari	0.364**	(0.163)	-0.390**	(0.163)
Alto Paraná	0.620***	(0.188)	0.148	(0.177)
Central	0.708***	(0.131)	-0.141	(0.181)
Neembucú	0.270	(0.168)	-0.414	(0.366)
Amambay	0.462***	(0.162)	-0.165	(0.155)
Canindeyú	0.175	(0.178)	0.235	(0.170)
Presidente Hayes	0.968***	(0.148)	-0.0409	(0.164)
Constant	2.605***	(0.880)	1.375**	(0.651)
Rho	0.1245	(0.163)	-0.3496	(0.2259)
Sigma	0.8085***	(0.0216)	0.7996***	(0.0257)
Lambda	0.101	(0.1322)	-0.2795†	(0.1814)
F	32.48		29.8	
Prob > F	0.0000		0.0000	
Observations	1916		1168	
Population Size	443893		320580	
Design df	495		217	
Reference Category	Until basic education, Formal Sector, Male, Agricultor Aborigen, Agriculture and services, Concepción			

Significance levels : † : 15% : ∗ : 10% ∗∗ : 5% ∗ ∗ ∗ : 1%

Table 14: Heckit Wage Reg. for rural stayers 1997-2005 (with control for migration to metropolitan and urban areas)

Variable	Coeff 97	Std.err	Coeff 05	Std.err
Secondary education	0.286***	(0.0730)	0.217***	(0.0824)
College education	0.525***	(0.147)	0.249*	(0.143)
Log of district household mean income at origin	0.339***	(0.0731)	0.604***	(0.0621)
Informal sector	-0.497***	(0.0592)	-0.285***	(0.0674)
Potential experience	0.0176**	(0.00742)	0.0221***	(0.00717)
Squared potential experience	-0.000366***	(0.000125)	-0.000433***	(0.000135)
Female	-0.442***	(0.0522)	-0.484***	(0.0605)
CEOs, Professionals, Technicians	0.616***	(0.139)	0.483**	(0.192)
Clercks, Operators, Office workers	0.453***	(0.0599)	0.326***	(0.0639)
Ethnic variable - Castellano	0.133	(0.134)	0.150	(0.113)
Public sector	0.299**	(0.125)	0.291	(0.191)
Electricity sector	0.0775	(0.357)	0.880**	(0.406)
Finance, commerce and Transport sector	0.176**	(0.0710)	0.136*	(0.0701)
San Pedro	-0.0421	(0.157)	-0.111	(0.167)
Cordillera	0.598***	(0.136)	-0.257	(0.161)
Guairá	0.392**	(0.174)	-0.517**	(0.245)
Caaguazú	0.0425	(0.148)	-0.225	(0.154)
Caazapá	0.400**	(0.169)	-0.541*	(0.278)
Itapúa	0.630***	(0.140)	-0.380**	(0.163)
Misiones	0.242	(0.192)	-0.501**	(0.249)
Paraguari	0.331**	(0.158)	-0.459***	(0.164)
Alto Paraná	0.622***	(0.186)	0.134	(0.182)
Central	0.700***	(0.128)	-0.186	(0.184)
Neembucú	0.279*	(0.167)	-0.460	(0.411)
Amambay	0.467***	(0.161)	-0.189	(0.159)
Canindeyú	0.121	(0.194)	0.145	(0.170)
Presidente Hayes	0.937***	(0.138)	-0.104	(0.169)
Constant	2.661***	(0.901)	1.458**	(0.663)
Rho	0.1111	(0.1166)	0.0564	(0.2374)
Sigma	0.8009***	(0.0221)	0.802***	(0.0256)
Lambda	0.0890	(0.0939)	0.0453	(0.1905)
F	31.14		30.13	
Prob > F	0.0000		0.0000	
Observations	1937		1201	
Population Size	448350		327757	
Design df	593		240	
Reference Category	Until basic education, Formal Sector, Male, Agricultor Aborigen, Agriculture and services, Concepción			

Significance levels : † : 15% : ∗ : 10% ∗∗ : 5% ∗ ∗ ∗ : 1%

Table 15: Heckit Wage Regressions for rural movers to metropolitan area 1997-2005 (with control for rural stayers)

Variable	Coeff 97	Std.err	Coeff 05	Std.err
Secondary education	-0.221	(0.163)	0.0548	(0.144)
Log of district household mean income at origin	0.0391	(0.135)	0.158	(0.160)
Informal sector	-0.198**	(0.0823)	-0.234	(0.159)
Potential experience	-0.0240*	(0.0139)	0.00310	(0.0119)
Squared potential experience	0.000641**	(0.000279)	-0.000146	(0.000184)
Female	-0.518***	(0.131)	-0.132	(0.177)
CEOs, Professionals, Technicians	1.730***	(0.234)	0.901***	(0.313)
Clercks, Operators, Office workers	0.00269	(0.0719)	-0.0748	(0.159)
Ethnic variable - Castellano	0.365***	(0.0922)	-0.0343	(0.126)
Public sector	-0.240	(0.194)	0.375**	(0.167)
Interior departments	0.0607	(0.0777)	-0.115	(0.157)
Border departments	0.317	(0.361)	-0.187	(0.228)
Constant	7.912***	(1.766)	6.897***	(1.851)
Rho	-0.4577***	(0.1936)	0.1424	(0.2889)
Sigma	0.3575***	(0.0415)	0.253	(0.04041)
Lambda	-0.1636**	(0.0808)	0.036	(0.077)
F	18.18		46.12	
Prob > F	0.000		0.0000	
Observations	1656		990	
Population Size	378397		271375	
Design df	502		223	
Reference Category	Until basic education, Formal Sector, Male, Agricultor Aborigen, Agriculture and services, Concepción			

Significance levels : † : 15% ∗ : 10% ∗∗ : 5% ∗ ∗ ∗ : 1%

Table 16: Heckit Wage Regressions for rural movers to metropolitan and urban areas 1997-2005 (with control for rural stayers)

Variable	Coeff 97	Std.err	Coeff 05	Std.err
Secondary education	-0.409	(0.378)	-0.0374	(0.138)
College education	1.208***	(0.369)	0.240767	(0.184)
Log of district household mean income at origin	0.256**	(0.109)	0.614***	(0.124)
Informal sector	-0.293**	(0.119)	-0.00404	(0.142)
Potential experience	0.00373	(0.0196)	0.00971	(0.0129)
Squared potential experience	-0.000124	(0.000409)	-0.000329	(0.000255)
Female	-0.295*	(0.169)	-0.405***	(0.143)
CEOs, Professionals, Technicians	0.255	(0.333)	0.287	(0.382)
Clercks, Operators, Office workers	-0.0486	(0.113)	-0.0155	(0.119)
Ethnic variable - Castellano	0.240**	(0.121)	0.117	(0.123)
Public sector	0.0349	(0.377)	0.337	(0.259)
Interior departments	-0.206*	(0.124)	0.0851	(0.115)
Border departments	0.0966	(0.165)	-0.166	(0.142)
Constant	5.110***	(1.412)	1.607	(1.411)
Rho	-0.4943	(0.2672)	-0.1346	(0.1812)
Sigma	0.6297***	(0.0877)	0.414***	(0.0369)
Lambda	-0.3113†	(0.1995)	-0.0558	(0.0759)
F	8.16		10.48	
Prob > F	0		0	
Observations	1728		1040	
Population Size	398712		283277	
Design df	556		255	
Reference Category	Until basic education, Formal Sector, Male, Agricultor Aborigen, Agriculture and services, Concepción			

Significance levels : † : 15% ∗ : 10% ∗∗ : 5% ∗ ∗ ∗ : 1%

Table 17: Probit structural forms, (probability to move to metropolitan area) 1997-2005

Variable	Coeff 97	Std.err	Coeff 05	Std.err
Wage differential (potential)	0.629***	(0.117)	0.692***	(0.171)
Married	-0.759***	(0.139)	-0.846***	(0.216)
Children	-0.266***	(0.0627)	-0.422***	(0.0804)
Average household education	0.154***	(0.0343)	0.185***	(0.0338)
Lagged Unemployment rate	0.0383***	(0.0125)	0.0168*	(0.00906)
Interior departments	0.335*	(0.183)	-0.701**	(0.324)
Border departments	-0.782**	(0.340)	-1.540***	(0.423)
Constant	-2.563***	(0.565)	-2.089***	(0.370)
F	15.37		11.55	
Prob > F	0.0000		0.0000	
Observations	1941		1185	
Population Size	452411		326530	
Design df	511		226	
Reference Category	Monoparental Household, Male Aborigen, Cordillera and the rest			

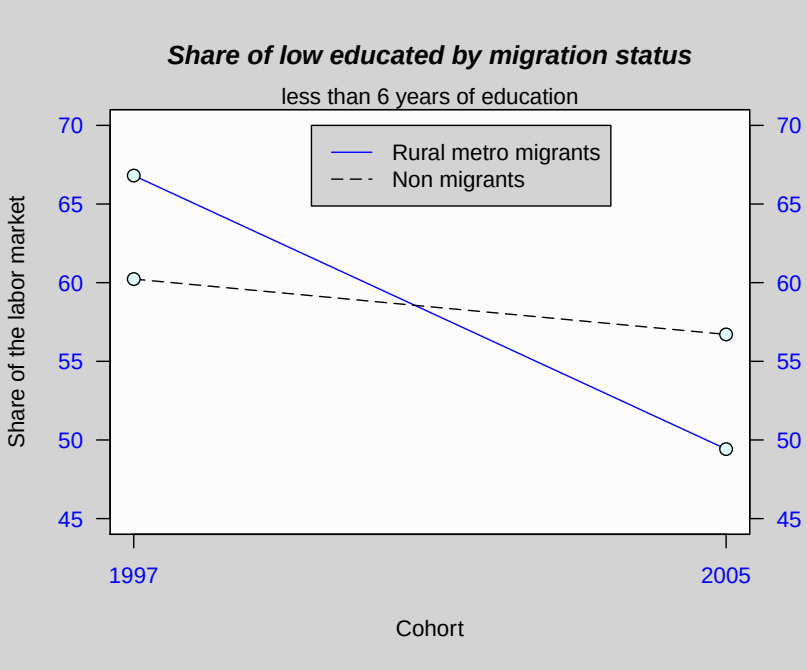
Significance levels : $*$: 10% $**$: 5% $***$: 1%

Table 18: Probit structural forms, (probability to move to metropolitan and urban areas) 1997-2005

Variable	Coeff 97	Std.err	Coeff 05	Std.err
Wage differential (potential)	0.186*	(0.101)	0.352*	(0.210)
Married	-0.612***	(0.113)	-0.546***	(0.147)
Children	-0.0771**	(0.0317)	-0.249***	(0.0630)
Average household education	0.133***	(0.0258)	0.131***	(0.0249)
Unemployment rate	0.0309***	(0.00968)	0.00631	(0.00616)
Interior departments	0.243	(0.147)	-0.288	(0.227)
Border departments	0.0428	(0.201)	-0.238	(0.192)
Constant	-2.526***	(0.450)	-1.637***	(0.282)
F	12.66		7.77	
Prob > F	0.0000		0.0000	
Observations	1981		1235	
Population Size	463215		338432	
Design df	565		258	
Reference Category	Monoparental Household, Male Aborigen, Cordillera and the rest			

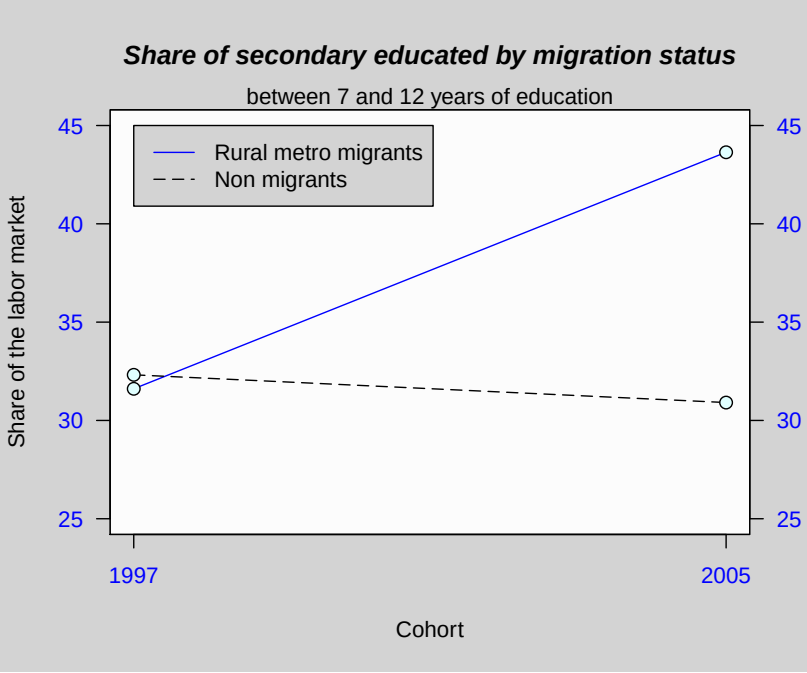
Significance levels : $*$: 10% $**$: 5% $***$: 1%

Figure 1: Share of low educated labor force by migration status



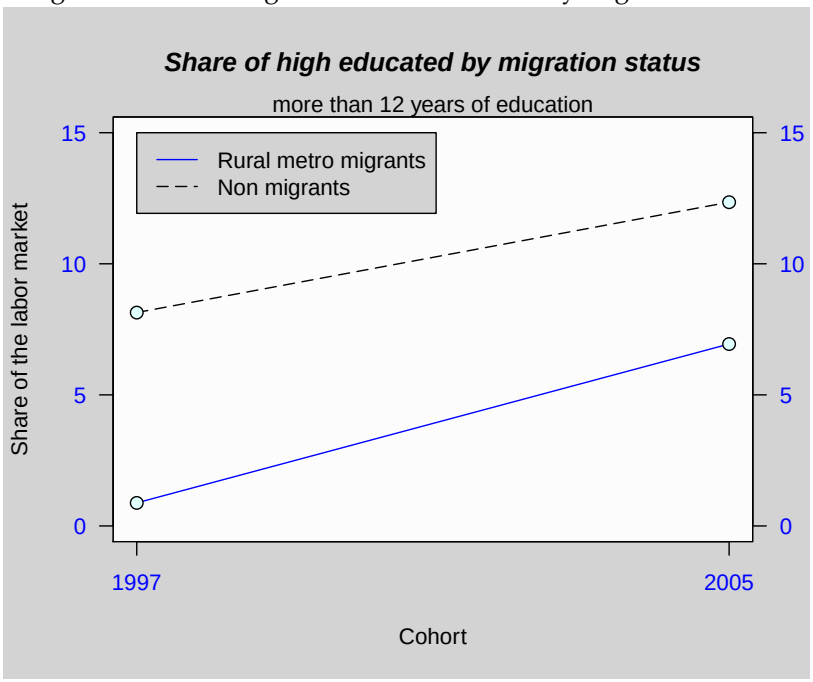
Source: Own elaboration, based on EIH 1997/98 and EPH 2005

Figure 2: Share of middle educated labor force by migration status



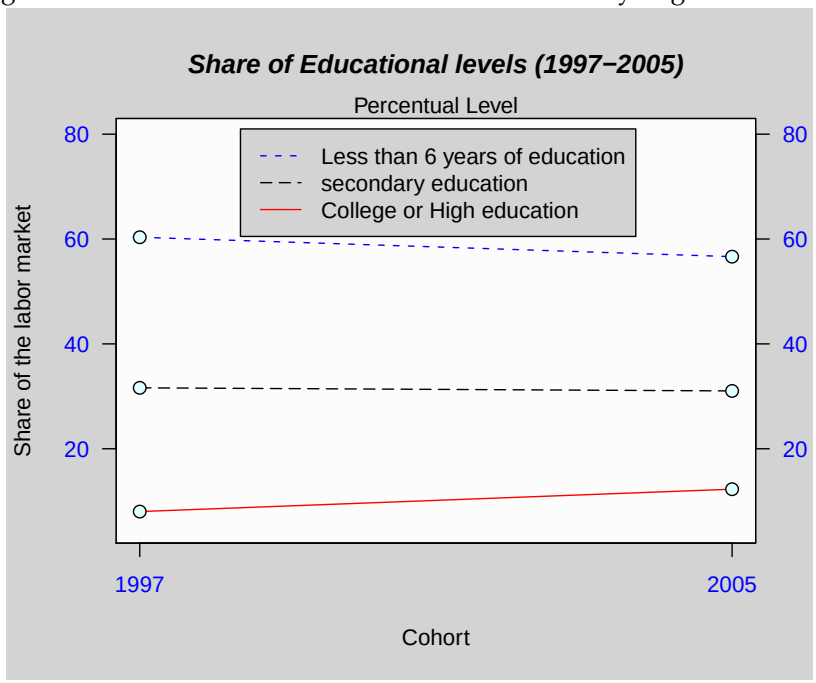
Source: Own elaboration, based on EIH 1997/98 and EPH 2005

Figure 3: Share of high educated labor force by migration status



Source: Own elaboration, based on EIH 1997/98 and EPH 2005

Figure 4: Share of educational level of the labor force by migration status



Source: Own elaboration, based on EIH 1997/98 and EPH 2005