A self selection analysis of employed rural-metropolitan migrants in Paraguay

Carlos Villalobos Barría*

Incomplete draft, February 28th, 2009

Abstract

Given the different income inequalities between rural, urban 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 labor market characteristics and moving cost, I use a switching regression model to evaluate if migrants from rural to metropolitan areas are somehow selected with respect to unobserved characteristics. I find evidence that rural metropolitan migrants are negatively selected but this effect also depends on the definition from rural, urban and metropolitan areas. Moreover, potential wage differentials, unemployment rates, household characteristics like marital status; the number of children in household and the average years of family education are always relevant and significant determinants of rural urban/metropolitan migration 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)

1 Introduction

Migration movements are substantial in Paraguay. Rural, urban and metropolitan areas work as both, source and destination areas. Because levels of human capital are considerably lower in rural areas than in urban and metropolitan ones, I have special interest in studying rural to urban/metro migration and how rural migrants integrate into urban/metro labor markets. Migration has consequences for households and regions and may work as a mechanism to equalize relative resource scarcities over regions. Individual migration decisions respond to economic opportunities as migrants seek higher returns to their attributes[1].

The decision to migrate is based on the perspective that the richest regions will be able to provide better levels of income or welfare to the migrants, relatively to what they would have in their place of origin. There are numerous factors that can explain the decision to migrate to a certain area. Among them are the relative indexes of standard of living, unemployment, household characteristics, and age and education attainment of the household head, etc. There are, however, some non-observable characteristics of the individuals, that influence the decision to migrate and also their level of earnings at the destination. The composition of the migrant population is a key issue to evaluate the consequences of migration. This composition concerns human capital characteristics and non-observable ability. For this reason, the study of self-selection of migrants plays an important role to understand the causes of rural-urban/metropolitan migration and their effects on labor markets, demand for public goods, public expenditure, investment, poverty and overall prospects of economic development.

Borjas (1987) used the Roy (1951) model to explain migration decision. In this model, the self selection of migrants is a function of the market return to their abilities and the income distributions at origin and destination. This model predicts that migrant will be negative selected if the inequality in the distribution of incomes is lower in the destination than in the origin. Since the inequality in earnings in Paraguay is higher in rural areas than in metropolitan and urban areas, the Roy model predicts that rural-urban/metropolitan migrants are negative selected.

[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). Do jobs follow people or do people follow jobs? This question of simultaneity resumes the phenomena property around the role of amenities on migration.

The empirical approach of this paper is a switching regression model presented by Goldfeld and Quandt (1973) with endogenous switching (Maddala, 1983). This empirical approach can be used as counterpart from a theoretical extended Roy model in the migration context, which allows for possible correlation between non-observable abilities and moving costs. Theoretically in this framework, the predictions of the standard Roy model can be even reversed.

2 Migration flows in Paraguay

I define a migrant as an individual who has moved to a different district in the past five years, excluding migrants from foreign countries. The EPH identifies the place of birth of migrants. Analyzing only recent migration and not lifetime migration has the drawback of not allowing the study of their assimilation at destination. Restricting the sample to a more homogeneous group makes the analysis of the migration decision more meaningful. I define Asuncion and some districts of Central as metropolitan area. All districts, which are not defined as rural areas, belong to urban areas.

While 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 brings 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 head of households 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[2].

Observing the data, I find that the rural-urban/metropolitan migration has a small size and became relatively less important respect to other kind of migration patterns like urban-urban, urban-metro, metro-urban, metro-metro, metro-rural, urban-rural migration. Although the low relative size of rural-urban/metropolitan migration, it represents a radical life-style change and therefore, I still interested to study this particular migration pattern. The lack of infrastructure, services and housing make this group especially vulnerable, because this group is characterized by lower incomes and education profiles than their urban or metropolitan counterparts.

[2] *Encuesta permanente de hogares* defines urban area as a locality formed by ten or more blocks, independently of the number of inhabitants. The Metropolitan area consists of Asuncion, Central´s districts *Capiatá, Fernando de la Mora, Guarambaré, Lambaré, Limpio, Mariano Roque Alonso, Nueva Italia, Nemby, San Antonio,*

Table 1: Migration flows, by origin and destination, 1997/98 (heads only)

	ORIGIN			
	Rural	Urban	Metropolitan	Total
Rural	23417	17050	7095	47562
% of dest. flow	49,23	35,85	14,92	100,00
% of origin flow	51,30	36,37	15,56	34,44
% of total flow	16,95	12,34	5,14	34,44
Urban	11267	16129	10332	37728
% of dest. flow	29,86	42,75	27,39	100,00
% of origin flow	24,68	34,40	22,67	27,32
% of total flow	8,16	11,68	7,48	27,32
Metropolitan	10961	13706	28157	52824
% of dest. flow	20,75	25,95	53,30	100,00
% of origin flow	24,01	29,23	61,77	38,25
% of total flow	7,94	9,92	20,39	38,25
Total	45645	46885	45584	138114
% of dest. flow	33,05	33,95	33,00	100,00
% of origin flow	100,00	100,00	100,00	100,00
% of total flow	33,05	33,95	33,00	100,00

Source: own elaboration, based on EIH 1997/98

In 1997-98, Table 1 shows that the main flow of migrating household heads comes from urban areas 33.35%, among them, 34.4% going to urban and 29.23% going to metropolitan destinations. The second main flow in 1997/98 is coming from rural areas and represents 33.05% of the total migration.

In 2005, Table 2, the main flow of migrating household heads comes from metropolitan areas and represents 41% of the total migration. Among them, 66.9% remain in other metropolitan areas, 10.95% go to urban areas and 22.14% go to rural areas.

In 1997 head of household leaving rural areas represent 33.05% of total migration. In 2005, this portion was 26.06%. Of these total, rural migrants going to urban or metro areas, represent 16.1% of the total migration in 1997/98 and 8.53% in 2005. The metropolitan origin migrants represent 33% of the total migrants in 1997/98. Among them, 84.4% goes to other urban or metropolitan areas and 15.6% goes to rural areas. Summarizing, rural movers to urban or metropolitan represent 16.1% of the total migration in 1997/98, while urban or metropolitan movers to rural represent as well 16.95% of the total.

San Lorenzo and some Central´s sub districts *Itá, Itaguá, Luque, Villa Elisa, Villeta.*

Table 2: Migration flows, by origin and destination, 2005 (heads only)

	ORIGIN			
	Rural	Urban	Metropolitan	Total
Rural	23928	20105	12403	56436
% of dest. flow	42,40	35,62	21,98	100,00
% of origin flow	67,29	44,82	22,14	41,36
% of total flow	17,54	14,74	9,09	41,36
Urban	6066	12342	6135	24543
% of dest. flow	24,72	50,29	25,00	100,00
% of origin flow	17,06	27,51	10,95	17,99
% of total flow	4,45	9,05	4,50	17,99
Metropolitan	5563	12415	37486	55464
% of dest. flow	10,03	22,38	67,59	100,00
% of origin flow	15,65	27,67	66,91	40,65
% of total flow	4,08	9,10	27,47	40,65
Total	35557	44862	56024	136443
% of dest. flow	26,06	32,88	41,06	100,00
% of origin flow	100,00	100,00	100,00	100,00
% of total flow	26,06	32,88	41,06	100,00

Source: own elaboration, based on EPH 2005

Concerning inequality, Table 3 reports the own calculated Gini coefficients for the hourly wage distributions for the studied years in rural, urban and metropolitan areas.

Table 3: Gini coefficients of hourly wages in Paraguay, by areas

Year	Rural	Urban	Metropolitan	Paraguay
1997	0.48539	0.51113	0.45244	0.50231
2005	0.52336	0.47467	0.49635	0.50981

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 grew slightly between 1997 and 2005. Rural and metropolitan areas became more unequal while the urban ones became more equal in the same time interval. The standard Roy model predicts for the Paraguayan migrants from rural areas to be negative selected to metropolitan areas in both years and positive selected to urban areas in 1997 and negative selected to the same destination in 2005.

3 The Data

For this study, the official data was provided by *Dirección General de Estadísticas, Encuestas y Censos (DGEEC) de Paraguay*. I use *Encuesta Permanente de Hogares 1992 and 2005 (EPH 1992, EPH 2005) and Encuesta Integral de Hogares 1997/98 and 2000/01 (EIH 1997/98, 2000/01)* . *EIH 1997/98 and EIH 2005* are a two-step sampling living standard survey which offers for each household the same probability of being sampled. Only the departments of Boquerón and Alto Paraguay are not represented in the survey. The sample size reach 5000 Households, with representativeness at country, urban, rural, strata and main departments (San Pedro, Itapúa, Caaguazú, Alto Paraná and Central) levels. *EPH 2005* contents almost the same variables as *Encuesta Integrada de Hogres (EIH) 1997/98*, but they are not 100% compatible[3].

According to the migrant definition, I distinguish two categories of individuals. *Movers* are the individuals who move from rural to metropolitan and urban/metropolitan areas anytime in the past five years and *Stayers* are the individuals who stay in rural areas longer as five years.

[3]To obtain a more detail see www. dgeec.gov.py

4 Theoretical framework

Migration is the result of a cost and 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 in the origin and destination, migrant with higher abilities tends 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 taking into account if moving costs decrease with the amounts of human capital. It is reasonable to assume, that the same characteristics which yields individuals to obtain higher wages can yield also to a reduction of the moving cost. If this is the case, is also plausible that individuals on the top of the income distribution at home (origin) decide to move to a host area with a more equal distribution of wages.

4.1 Extended Roy model

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) 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.

Assume now, that C represent 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. Migration occurs if $\frac{W_2 - W_1}{W_1} > C$, which is approximately $lnW_2 - lnW_1 > C$

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

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 $I = (u_1 - u_2 - \gamma + e_2 - e_1 - \varepsilon) > 0$ and stay if $I \leq 0$. Assuming the normality conditions and defining $\sigma^{\nu} = \sqrt{Var(e_2 - e_1 - \varepsilon)}$, $z = -\frac{u_2 - u_1 - \gamma}{\sigma^{\nu}}$ and $\eta = \frac{e_2 - e_1 - \varepsilon}{\sigma^{\nu}}$, 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

$$E(lnW_1 | I > 0) = u_1 + \sigma_{e1\eta}\lambda(z) \tag{5}$$

and the observed wages at the destination as

$$E(lnW_2 | I > 0) = u_2 + \sigma_{e2\eta}\lambda(z) \tag{6}$$

where $\sigma_{e1,2\eta}$ represent the covariance between unobservable abilities and the normalized net abilities considering the moving costs disturbances between home and host regions. $\lambda(Z) = \frac{\phi(z)}{1 - \Phi(z)}$ is the inverse of Mill´s ratio and $\phi()$ the standard normal density function. Because $\lambda(z)$ is strictly positive, on average, a migrant is better off than an average person in the home region if $\sigma_{e1\eta} > 0$, in this case the mover is positively selected and better off than an average person in the host region if $\sigma_{e2\eta} > 0$. If $\sigma_{e1\eta} < 0$, the migrants are negatively selected or, in other words, comes from the bottom tail of the wages distribution in the home region.

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 (Maddala and Nelson, 1975; Maddala, 1983). Equation (1) and (2) can be rewritten as

$$lnW_{1i} = X_{1i}\beta_1 + \mu_{1i} \tag{7}$$

$$lnW_{2i} = X_{2i}\beta_2 + \mu_{2i} \tag{8}$$

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

$$I_i = \delta(lnW_{2i} - lnW_{1i}) - Z_i\psi - \varepsilon_i \tag{9}$$

where additionally the migration cost $Z_i\psi + \varepsilon_i = C$ (counterpart from (3)) depends again on personal characteristics in Z_i and unobservable ε_i. In order to identifier 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 tree steps 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.

$$I_i^* = \delta(X_{2i}\beta_2 - X_{1i}\beta_1) - Z_i\psi + \delta(\mu_{2i} - \mu_{1i}) - \varepsilon_i = Z_i^*\psi^* + \varepsilon_i^* \tag{10}$$

The parameter vector $\hat{\psi}^*$ can be suitable estimated and therefore the inverse Mills' ratio for stayers and migrants can be calculated. Equations (7) and (8) can be estimated for each individual and as consequence a potential differential wage variable can be also calculated. Both equations can be estimated for stayers and movers with control for selection bias. For stayers

$$lnW_i = X_i\beta_1 - \sigma_{e1\eta}\frac{\phi(Z_i^*\hat{\psi}^*)}{\Phi(Z_i^*\hat{\psi}^*)} + \mu_{1i} \tag{11}$$

and for movers

$$lnW_i = X_i\beta_2 - \sigma_{e2\eta}\frac{\phi(Z_i^*\hat{\psi}^*)}{1-\Phi(Z_i^*\hat{\psi}^*)} + \mu_{2i} \tag{12}$$

In this paper I use the Heckman Maximum Likelihood estimator for survey data which also take into account the correlation between primary sample units avoiding the underestimation of standard errors and as consequence avoiding the overestimation of sample selection bias or possible self-selectivity of migrants.

The last step to estimate structurally the index function is by using as regressors the estimated potential wage differential and other variables which determines migration probability but not through wages. I used unemployment rates, average family education and the number of children in household which also as were the excluded variables in (11) and (12) to identify (10).

5 Estimation Results

First, the strategy consist to obtain consistent estimates of the individual probability to migrate, this estimates yield to obtain unbiased estimates for wages in Rural, Metropolitan and Urban/Metropolitan areas. Lastly the estimated potential wage differential, which is not more endogenous, is included as regressor in the structural probit for migration. Since in Paraguay no panel data with labor market information is available, forthe probit models, I used lagged values of the interest explanatory variables which were obtained from available past survey whit similar characteristics, 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 2005.*

5.1 Probit, reduced form

Following Harris and Todaro (1970) the individual 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, and $C(0)$is the migration cost. Beyond the potential differential wage, personal characteristics such as education, potential experience, sex and ethnic origin joint with household composition variables such as marital status, children in household are expected to influence the probability to move. Lagged departmental unemployment rates at origin could impact the probability to be employed and in this way affect expected earnings.

The grade 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 from the monetary contribution of each member. The cost in case of migration will be proportional to the household welfare dependence to any household member and therefore, for example, an individual would be able to move if he knows that in case of migration, the household will be not fall into poverty.

I introduce this cost proposition by using an indicator for the household income capacity generation. I used the average family education in years.

Tables 1 and 2 in the appendix show the probit reduced forms for the probability to migrate to metropolitan area and to urban/metropolitan areas in 1997 and 2005.

The probit regressions indicate a negative relationship between education grade and the probability to move. This effect is stronger for individuals with collage education but can also reflect that those people are not willing to migrate because they are permanently employed. Being married is associated with a reduction in the probability to migrate, but this effect appears to be significant only in 1997. Ethnic variable called *Castellano* has a strong and significant impact pro migration. This variable is a design variable being equal to the unity when the individual speaks exclusively *Castellano* and therefore is supposed does not have an aboriginal origin. Lagged departmental unemployment rates at origin appear to stimulate migration. Potential experience appears to have a significant non linear impact on migration probability. The mentioned output controls for interior and border departments.

5.2 Heckit wage regressions

Tables 3, 4, 5 and 6 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 variable excludes in the wage regression, in order to identify the system, are the number of children in household, marital status and the average family education in years. These variables should not be correlated with the residual of the wage regressions.

The log hourly wages are explained by human capital accumulation indicators such as education grades and potential experience. Those 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 control also for public sector and department where the individual is employed.

Based on the probit reduced form specification, I can examine if the second terms in equations (11) and (12) are significant positively or negatively different from zero. The sign of the Lambdas indicates if the migrant/stayers are positively or negatively selected. I find strong evidence that in 1997 rural metropolitan movers were negative selected while to urban/metropolitan areas this evidence is weaker. For 2005 the evidence concerning the self selectivity of movers is ambiguous and insignificant.

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 7 and 8 also include as regressors marital status, average household education, lagged district unemployment rates, and controls for geographic location.

The results are highly robust and almost all variables for both years have expected signs and statistically significance. The potential wage differential appears to be an important determinant of the probability to move. In the extreme case, move to metropolitan area, the wage differential appears to have a higher impact on this probability than to move to urban/metropolitan areas. Marital status appears to recover their theoretical importance determining probability of migration.

Average household education appears to incentive migration according to the idea that the cost in case of migrate will be proportional to the household welfare dependence to any household member. For this reason the higher the average household education the lower the dependence on an individual household member and therefore higher the probability to move for this individual.

As expected, the higher the lagged departmental unemployment rates, the higher the propensity to move. This effect is reported to be higher in 1997 than for 2005.

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 cost could be correlated with abilities in labor markets. To investigate empirically this issue, I used a switching regression model with endogenous switching.

The results show strong evidence that migrants were negatively selected in 1997 to metropolitan areas. For 2005, I found no evidence about self selectivity of migrants. The findings of this study is that rural metropolitan and rural urban/metropolitan migration do not represent a lost of the relative human capital stock for rural areas in relation with the metropolitan or urban/metropolitan ones. For 2005, rural movers to metropolitan or urban/metropolitan areas appear to be a random sample of the Paraguayan rural labor market.

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 in household and the average family education.

I find evidence that the potential wage differential determines the probability to move. The higher is the wage differential, the higher is the probability to move. Marital status and children in households also affect the migration probability in the expected ways as costs factors.

The average family education is an important determinant of the probability of migration. The grade in 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 from the potential monetary contribution of each member to the household. The cost in case of migrate will be somehow proportional to the household welfare dependence to any household member and therefore, is not surprising that higher levels of average education implies higher degree of freedom to the individuals by reducing the costs in case of migration.

7 References

Borjas, G. (1999): "Self-Selection and the Earnings of Immigrants", in *Labor Economics* ed. by O. Ashenfelter, pp.305-335, Princeton University.

Brücker, Herbert and Trübswetter, Parvati, "Do the Best Go West? An Analysis of the Self-Selection of Employed East-West Migrants in Germany" (January 2004). IZA Discussion Paper No. 986. Available at SSRN: http://ssrn.com/abstract=494142

Buchinsky, M. (1998): "The Dynamics of Changes in the Female Wage Distribution in the USA: A Quantile Regression Approach", *Journal of Applied Econometrics*, 13, 1-30.

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

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". In Handbook of Families and Population Economics. Rosenzweig and Stark (Eds.). North-Holland, Amsterdam.

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.

Ichimura, H., (1993). "Semiparametric least-squares (SLS) and weighted SLS estimation of single-index models". *Journal of Econometrics,* Vol. 58, 71-120.

Koenker, R e G. Bassett (1978): "Regression Quantiles", *Econometrica*, 46, pp. 33-50.

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

Machado, J. and Koenker, R. (1999). "Goodness of fit and related inference processes for quantile regression", *Journal of the American Statistical Association*, Vol. 94, pp.1296-1310.

Mincer, J. (1994): "Human Capital: a Review", in 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.

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

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", in Migration and Labor Market Adjustment. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Table 1: Probit reduced forms, probability to move to metropolitan area 1997-2005

Variable	Coeff 97	Std.err	Coeff 05	Std.err
Basic and Secondary education	-0.611**	(0.238)	-0.338	(0.249)
Collage 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	No education, Monoparental Household, Male Aboriginal origin, Cordillera and the rest			

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

Table 2: Probit reduced forms, probability to move to metropolitan and urban areas 1997-2005

Variable	Coeff 97	Std.err	Coeff 05	Std.err
Basic and Secondary education	-0.144	(0.133)	-0.268	(0.170)
Collage 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	1980		1235	
Population Size	464572		338432	
Design df	567		258	
Reference Category	No education, Monoparental Household, Male Aboriginal origin, Cordillera and the rest			

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

Table 3: Heckit Wage Reg. for rural stayers 1997-2005 with control for migration to metropolitan area

Variable	Coeff 97	Std.err	Coeff 05	Std.err
Basic and Secondary education	0.210***	(0.0755)	0.223***	(0.0813)
Collage education	0.529***	(0.164)	0.308**	(0.142)
Log of dist. mean household inc.at origin	0.329***	(0.0711)	0.610***	(0.0605)
Informal sector	-0.439***	(0.0632)	-0.268***	(0.0674)
Potential experience	0.0170***	(0.00654)	0.0231***	(0.00699)
Squared potential experience	-0.000366***	(0.000112)	-0.000381***	(0.000134)
Female	-0.406***	(0.0495)	-0.490***	(0.0602)
CEOs, Professionals, Technicians	0.720***	(0.133)	0.457**	(0.187)
Clercks, Operators, Office workers	0.433***	(0.0608)	0.328***	(0.0632)
Ethnic variable - Castellano	0.105	(0.141)	0.171	(0.112)
Public sector	0.251*	(0.130)	0.309	(0.188)
Electricity sector	0.0261	(0.351)	0.846**	(0.390)
Finance, commerce and Transport sector	0.185***	(0.0687)	0.121*	(0.0687)
San Pedro	-0.0328	(0.157)	-0.0892	(0.166)
Cordillera	0.537***	(0.139)	-0.218	(0.161)
Guairá	0.359**	(0.180)	-0.436*	(0.239)
Caaguazú	0.0310	(0.150)	-0.189	(0.154)
Caazapá	0.372**	(0.166)	-0.458*	(0.277)
Itapúa	0.596***	(0.141)	-0.346**	(0.162)
Misiones	0.0514	(0.223)	-0.450*	(0.236)
Paraguari	0.354**	(0.163)	-0.390**	(0.163)
Alto Paraná	0.612***	(0.186)	0.148	(0.177)
Central	0.690***	(0.131)	-0.141	(0.181)
Neembucú	0.264	(0.164)	-0.414	(0.366)
Amambay	0.457***	(0.163)	-0.165	(0.155)
Canindeyú	0.181	(0.180)	0.235	(0.170)
Presidente Hayes	0.951***	(0.150)	-0.0409	(0.164)
Constant	2.755***	(0.880)	1.375**	(0.651)
Rho	0.078	(0.1908)	-0.349†	(0.225)
Sigma	0.806***	(0.022)	0.799	(0.026)
Lambda	0.063	(0.154)	-0.279†	(0.181)
F	33.33		29.80	
Prob > F	0.0000		0.000	
Observations	1916		1168	
Population Size	443893		320580	
Design df	495		217	
Reference Category	No education, Formal Sector, Male, Agricultor Aboriginal origin, Agriculture and services, Concepción			

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

Table 4: 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
Basic and Secondary education	0.273***	(0.0771)	0.217***	(0.0824)
Collage education	0.643***	(0.163)	0.249*	(0.143)
Log of dist. mean household inc.at origin	0.338***	(0.0728)	0.604***	(0.0621)
Informal sector	-0.470***	(0.0629)	-0.285***	(0.0674)
Potential experience	0.0130*	(0.00730)	0.0221***	(0.00717)
Squared potential experience	-0.000300**	(0.000122)	-0.000433***	(0.000135)
Female	-0.436***	(0.0505)	-0.484***	(0.0605)
CEOs, Professionals, Technicians	0.646***	(0.137)	0.483**	(0.192)
Clercks, Operators, Office workers	0.427***	(0.0608)	0.326***	(0.0639)
Ethnic variable - Castellano	0.126	(0.135)	0.150	(0.113)
Public sector	0.214*	(0.126)	0.291	(0.191)
Electricity sector	0.0437	(0.338)	0.880**	(0.406)
Finance, commerce and Transport sector	0.193***	(0.0698)	0.136*	(0.0701)
San Pedro	-0.0428	(0.158)	-0.111	(0.167)
Cordillera	0.585***	(0.135)	-0.257	(0.161)
Guairá	0.380**	(0.173)	-0.517**	(0.245)
Caaguazú	0.0309	(0.148)	-0.225	(0.154)
Caazapá	0.387**	(0.165)	-0.541*	(0.278)
Itapúa	0.622***	(0.140)	-0.380**	(0.163)
Misiones	0.231	(0.191)	-0.501**	(0.249)
Paraguari	0.338**	(0.159)	-0.459***	(0.164)
Alto Paraná	0.606***	(0.183)	0.134	(0.182)
Central	0.678***	(0.128)	-0.186	(0.184)
Neembucú	0.269*	(0.163)	-0.460	(0.411)
Amambay	0.454***	(0.162)	-0.189	(0.159)
Canindeyú	0.121	(0.196)	0.145	(0.170)
Presidente Hayes	0.944***	(0.147)	-0.104	(0.169)
Constant	2.711***	(0.902)	1.458**	(0.663)
Rho	0.119	(0.113)	0.056	(0.237)
Sigma	0.802***	(0.023)	0.802***	(0.026)
Lambda	0.095	(0.0904)	0.045	(0.191)
F	34.59		30.13	
Prob > F	0.0000		0.0000	
Observations	1935		1201	
Population Size	448935		327757	
Design df	541		240	
Reference Category	No education, Formal Sector, Male, Agricultor Aboriginal origin, Agriculture and services, Concepción			

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

Table 5: 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
Basic and Secondary education	-0.298*	(0.164)	0.0486	(0.105)
Log of dist. mean household inc.at origin	0.116	(0.217)	0.175	(0.144)
Informal sector	-0.192**	(0.0775)	-0.179	(0.187)
Potential experience	-0.0273**	(0.0132)	0.00637	(0.0117)
Squared potential experience	0.000718***	(0.000266)	-0.000189	(0.000215)
Female	-0.559***	(0.134)	-0.129	(0.186)
CEOs, Professionals, Technicians	1.933***	(0.254)	0.976***	(0.315)
Clercks, Operators, Office workers	-0.0115	(0.0763)	-0.0115	(0.161)
Ethnic variable - Castellano	0.297***	(0.100)	-0.0452	(0.123)
Public sector	-0.241	(0.183)	0.288*	(0.158)
Finance, commerce and Transport sector	0.0487	(0.0833)	0.0419	(0.147)
Central	0.0812	(0.136)	-0.0975	(0.119)
Constant	6.898**	(2.919)	6.635***	(1.732)
Rho	-0.506**	(0.181)	0.085	(0.216)
Sigma	0.356***	(0.039)	0.252***	(0.042)
Lambda	-0.180**	(0.077)	0.021	(0.057)
F	16.07		73.96	
Prob > F	0.0000		0.0000	
Observations	1656		990	
Population Size	378397		271375	
Design df	502		223	
Reference Category	No education, Formal Sector, Male, Agricultor Aboriginal origin, Agriculture and services, Concepción			

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

Table 6: 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
Basic and Secondary education	-0.129	(0.288)	0.0118	(0.128)
Collage education	0.272	(0.340)	0.0723	(0.181)
Log of dist. mean household inc.at origin	-0.118	(0.120)	0.617***	(0.144)
Informal sector	-0.249**	(0.0999)	0.192	(0.160)
Potential experience	-0.0119	(0.0171)	-0.00613	(0.0142)
Squared potential experience	0.000101	(0.000359)	-0.0000488	(0.000275)
Female	-0.164	(0.139)	-0.460***	(0.136)
CEOs, Professionals, Technicians	0.811**	(0.388)	0.501**	(0.242)
Clercks, Operators, Office workers	0.0347	(0.121)	0.0919	(0.108)
Ethnic variable - Castellano	0.311***	(0.119)	0.0307	(0.130)
Public sector	-0.351	(0.336)	0.457**	(0.184)
Finance, commerce and Transport sector	0.294**	(0.123)	0.0931	(0.120)
Central	0.0820	(0.106)	-0.120	(0.161)
Constant	9.620***	(1.586)	1.539	(1.636)
Rho	-0.1609	(0.313)	-0.1109	(0.223)
Sigma	0.559***	(0.042)	0.360***	(0.024)
Lambda	-0.090	(0.176)	-0.0399	(0.081)
F	104.13		754.93	
Prob > F	0.0000		0.0000	
Observations	1734		1040	
Population Size	403234		283277	
Design df	558		255	
Reference Category	No education, Formal Sector, Male, Agricultor Aboriginal origin, Agriculture and services, Concepción			

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

Table 7: Probit structural forms, probability to move to metropolitan area 1997-2005

Variable	Coeff 97	Std.err	Coeff 05	Std.err
Wage differential (potential)	0.551***	(0.110)	0.794***	(0.188)
Married	-0.724***	(0.139)	-0.856***	(0.217)
Children	-0.270***	(0.0620)	-0.421***	(0.0811)
Average household education	0.165***	(0.0336)	0.181***	(0.0341)
Lagged Unemployment rate	0.0364***	(0.0124)	0.0179*	(0.00917)
Interior departments	0.246	(0.176)	-0.622*	(0.329)
Border departments	-0.987***	(0.339)	-1.451***	(0.421)
Constant	-2.573***	(0.568)	-2.057***	(0.370)
F	13.87		11.55	
Prob > F	0.0000		0.0000	
Observations	1941		1185	
Population Size	452411		326530	
Design df	511		226	
Reference Category	Monoparental Household, Cordillera and the rest			

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

Table 8: 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)	0.375***	(0.106)	0.306*	(0.178)
Married	-0.600***	(0.115)	-0.541***	(0.146)
Children	-0.0632**	(0.0291)	-0.254***	(0.0633)
Average household education	0.127***	(0.0263)	0.133***	(0.0247)
Lagged Unemployment rate	0.0305***	(0.00987)	0.00571	(0.00622)
Interior departments	0.360**	(0.167)	-0.247	(0.235)
Border departments	-0.0130	(0.201)	-0.184	(0.185)
Constant	-2.476***	(0.470)	-1.661***	(0.272)
F	12.27		7.78	
Prob > F	0.0000		0.0000	
Observations	1980		1235	
Population Size	464572		338432	
Design df	567		258	
Reference Category	Monoparental Household, Cordillera and the rest			

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