Taller Nacional sobre *"Migración interna y desarrollo en Brasil: diagnóstico, perspectivas y políticas"*

30 de Abril 2007, Brasilia, Brasil

Organizado por la Comisión Económica para América Latina y el Caribe, CELADE-División de Población, con el apoyo y auspicio del Banco Interamericano de Desarrollo (BID)

The selectivity of migration in Brazil: implications for rural poverty

André Braz Golgher

The selectivity of migration in Brazil: implications for rural poverty[1]

André Braz Golgher[2]

ABSTRACT

Poverty levels in Brazil present a remarkable spatial heterogeneity. Rural parts of the Northeast and the North regions present the greatest proportions of poor people. Migration from and to rural areas may have an impact on regional poverty levels, in particular due to the selectivity of migration. Mathematical simulations, based on the Roy and the human capital models, discussed this selectivity, and multinomial logistic models analyzed empirically some aspects of these simulations. It was observed a general tendency of negative selection in rural/rural flows and a positive one in rural/urban and longer steps of migration.

Key words: migration, selectivity, poverty, Brazil, Latin America.

1. Introduction

Despite the Brazilian general socioeconomic development in recent years, for instance the observed increase in schooling levels (Riani et al, 2004); poverty and income inequality did not present this same tendency. Between 1977 and 1999, indicators related to these two variables showed stable values with short term fluctuations (Barros et al, 2000), and just very recently that it was verified a slight advance on them (IBRE/FGV, 2005).

Hoffmann (2000) showed that poverty levels in Brazil present a remarkable spatial heterogeneity. Among the macroregions of this country, the Northeast Region had only 29% of the Brazilian population and 53% of the poor people in 1997. Ferreira et al (2000) estimated the proportion of poor people for different regions in Brazil in 1996. They verified that the Northeast Region had the greatest numbers, specially the rural parts, and was followed by the North Region. In the other macroregions of Brazil, Southeast, South and Center-West, the numbers were smaller, but still quite expressive.

There are many phenomena that may have an impact on regional and individual poverty levels and migration from and to rural areas is one of them. The influence of migration on regional poverty depends on the magnitude of the flows and also on their composition, which is related to the selectivity of migration. In this paper the relationship between poverty and migration selectivity is examined. In particular, it discusses some of the individual characteristics that might impact in the probability that a person is a migrant, similarly as was done in Sandefur et al (1991).

In order to do so, this work was divided in seven sections including this introduction. In the next, the human capital model is briefly presented, where the selectivity of the migratory process, that

[1] This material is based in works done with support and contribution of The United States Agency for International Development (AID), and with a subcontract of Broadening Access and Strengthening Input Market Systems (BASIS) / Collaborative Research Support Program (CRSP)/University of Wisconsin – Madison conferred for the Regents of the University of California, Riverside. Opinions, commentaries, conclusions and recommendations are of the exclusive responsibility of the authors and not necessary the ones of the Regents of the University of California, BASIS/CRSP and/or AID.
[2] Rua Cristina, 1092/302. BH/MG Brazil 30330-130. Visiting professor in Cedeplar/UFMG (031) 91419680. agolgher@cedeplar.ufmg.br

is the central point of the empirical analyses, is also discussed. As we will see in the following section, which shows some descriptive data, the migrant is not a random sample of the population and may have different characteristics when compared to non-migrants. Then, in the fourth section, based on these previous discussions, shows a mathematical simulation that discussed the selectivity of migration. Section five presents the methodology used in the empirical analysis, which were done with a multinomial logistic model applied to micro data. In section six, the main empirical results are showed, including some simulations. The last section presents the conclusions.

2. Human Capital Model and the Selectivity of Migration

Many models that analyze the relation between migration, income and poverty utilize the human capital model of migration as the theoretical foundation. The model assumes a rational individual migrates if the expected net return of migration is positive, and if so, he/she maximizes his/her utility among the possible destinies (Stillwell and Congdon 1991). The equation below presents this relation:

$$G_{ij} = \int_0^t (U_{jt} - U_{it}) e^{-pt} dt - C_{ij} > 0,$$

where G_{ij} is the net return of migration between localities **i** and **j**; U_{jt} is the utility (or expected wages) in **j**, which is a possible destiny of the migrant in time **t**; U_{it} is the utility of the person in the currently origin **i** in time **t**; ρ is the discount rate; and C_{ij} are the costs of migration between **i** and **j**.

Factors that influence the utility of individuals include personal attributes (sex, age, income, schooling, etc.), regional characteristics (unemployment rate, per capita income, climate, criminality, housing costs, leisure possibilities, etc.) and the interaction between these variables (Stillwell and Congdon 1991).

The costs of migration are normally a function of the distance between the origin and the destiny of the migrant. These costs can be monetary, psychological, of opportunity, of adaptation, etc (Stillweel and Congdon, 1991). It is believed that the costs are an increasing concave function of distance (Cadwallader, 1992; Bell et al, 1990). If the person cannot cope with the cost of a long step migration, an alternative would be to migrate in two or more shorter steps of migration with lower costs associated to them **(**Cadwallader, 1992).

Besides the distance, many other factors influence the costs of migration. Among them, and of particular importance for this study, is the presence of effective social nets between the potential migrants and persons in the destiny that may diminish decisively these costs by a series of reasons, enhancing the chance of migration, or even making the change of place of residence possible (Duarte, 1979; Gugler, 1992; Hollnsteiner-Racelis, 1988; Massey et al, 1993; Todaro, 1980).

A type of migration that presents lower costs than indicated by the distance is the return one, when the individual migrates and then returns to his/her original place of residence. This type of migration can be induced by many factors: poor evaluation or lack of adaptation in the destiny; regional changes occurred after migration (for instance, soil depletion in the destiny); life cycles (for instance, retirement), etc.

However, it must be emphasized that, as migration implicate in monetary and other types of costs, the individual must hold a minimum amount of capital to have migration as an option. This may not be a feasible situation for all the population strata. Poor people, specially the chronic or extremely poor ones, may not have this possibility (Kothari 2002).

Therefore, migrants are not a random sample of the population; on the contrary, there is a selection of those with some specific characteristics that will preferentially migrate to particular places. Many authors have discussed the differences between migrants and non-migrants. Borjas (1996) showed that young individuals with higher levels of formal education tended to be more mobile. Greenwood (1985) observed that household heads that were unemployed had a greater propensity to migration. Sandefur et al (1991) verified that Whites were less prone to be trapped in places with shrinking opportunities in the labor market, when compared to other ethnic groups.

Generally, it is believed that a typical migrant is a young adult, bachelor, with a reasonable level of formal education, with more effective social nets and that is more labor market oriented (Castiglione, 1989). However, what a typical migrant actually is depends on the context being analyzed and the type of migration that is being studied (De Haan, 1999; Todaro, 1980). For instance, sex ratios of migrant flows are directly influenced by the supply and demand sides of the labor market; typically rural areas may relatively better absorb men, while the contrary may occur in urban areas.

Specifically in Brazil, females were the majority for non-migrants and for migrants, although the predominance of women was slightly greater in this last group. However, some larger differences were observed between states. Some states had proportions of men above the Brazilian mean and the migrant flows were increasing these proportions. Consequently, although females were the majority of the migrants in Brazil, men outnumbered women in many states. Some differences regarding age distribution of the population were also noticed. As expected by the description of the typical migrant, migrants show greater proportions of intermediate age groups, while non-migrant are overrepresented by children and elderly. Concerning schooling and income, migrants showed larger values for both variables when compared with non-migrants for all states in Brazil. These show that there is general trend of positive selection (Golgher, 2006a).

As mentioned in the introduction of this paper, Brazil is regionally unequal. This spatial heterogeneity influences the flows of migrants, both quantitatively and qualitatively. Hence, because of factors, such as, spatial localization of the origin and of the destiny, type of flow, distance of migration, etc., the flows may present remarkable differences in many aspects. Next section shows some specific descriptive data about the differences between migrants and non-migrants in Brazil.

3. Descriptive Data

This section presents some descriptive data about the selectivity of migration in Brazil and also introduces some topics that will be analyzed empirically, such as the differences observed for distinct types of migration. The information was obtained with the use of the Brazilian Demographic Census of 2000. This database has the information in where the person lived five years before the Census research and the current place of residence. Individuals that declared different municipalities were considered migrants in the period of 1995-2000 (see Carvalho et al, 1992, and Rigotti, 1999, for a methodological discussion about migratory data in Brazilian Census).

In this paper, rural non-migrants and migrants that had as origin the same rural area are compared in the empirical analyses. Some data about them will be presented in the next tables. Table 1 shows the proportions of non-migrants and of migrants of different types for individuals with age between 18 and 64 years. Notice that the majority of them, around 80%, were non-migrants. In the Center-West, a region with many rural areas of population attraction, such as the north of Mato Grosso and the area around the Federal District, the value is slightly under 75%. As can also be seen in the table, most of the migrants are intrastate ones in all regions and the rural/rural ones are a little less numerous than the correspondent rural/urban. Only around 4% of the individuals were interstate migrants and approximately half of then were between states that are neighbors, indicating that the great majority of the migrants migrated in short steps, as expected by the human capital model.

Table 1. Proportion of migrants of different types and rural non-migrants

Type of flow	Proportion of migrants of different types and rural non-migrants by macroregion				
	North Region	Northeast Region	Southeast Region	South Region	Center-West Region
Rural/urban intrastate	7.1	5.8	9.0	9.3	10.3
Rural/urban between neighbors	2.0	1.1	2.0	1.2	3.0
Rural/urban between non-neighbors	1.1	2.2	0.5	1.5	1.7
Rural/rural intrastate	5.6	4.0	5.0	6.0	8.5
Rural/rural between neighbors	1.2	0.7	0.8	0.5	1.9
Rural/rural between non-neighbors	0.3	0.4	0.3	0.8	0.9
Rural non-migrants	82.6	85.7	82.4	80.6	73.7
Total	1944790	7918930	4197819	3104775	957544

Source: FIBGE, 2000.

Table 2 shows data about the proportion of men for each type of flow and compared this data with rural non-migrants. As discussed above, women are the majority in Brazil for both migrants and non-migrants. However, this is not true for non-migrants for rural Brazil, as can be seen in the last line of the table, in all macroregions. For rural/rural flows this also occurs for all flows. Rural areas are relatively more attractive for men than for women. Women were the majority in four short step rural/urban flows, especially in the Northeast Region, which had the lowest proportion of men for most of the flows.

Table 2. Comparison between flows of migrants of different types and rural non-migrants for proportion of men

Type of flow	Proportion of men in flows of migrants of different types and for rural non-migrants				
	North Region	Northeast Region	Southeast Region	South Region	Center-West Region
Rural/urban intrastate	49	46	49	50	52
Rural/urban between neighbors	51	48	50	52	51
Rural/urban between non-neighbors	52	50	51	53	55
Rural/rural intrastate	56	51	54	52	58
Rural/rural between neighbors	54	53	56	53	58
Rural/rural between non-neighbors	54	55	57	57	57
Rural non-migrants	54	52	53	52	55

Source: FIBGE, 2000.

In table 3, the flows of migrants are compared to non-migrants for the proportion of White/Asiatic. The first point to emphasize is that the South Region has much greater proportions than any other for this ethnic group, and that the North and Northeast have the highest proportions of Blacks/Pardos/Indigenous. For these last two regions, short distance flows have greater proportions of Blacks/Pardos/Indigenous than longer distance ones, indicating that this last group preferentially migrates in low cost steps. In the Southeast and the South regions, occurred the contrary; the long distance steps of migration had lower values for the proportion of White/Asiatic, partly due to the return migration to the Northeast and North regions that have much lower costs than expected by the distance owing to effective social nets. The Center-West Region was somewhat similar to the first two discussed.

Table 3. Comparison between flows of migrants of different types and rural non-migrants for proportion of White/Asiatic

Type of flow	Proportion of White/Asiatic for flows of migrants of different types and for rural non-migrants				
	North Region	Northeast Region	Southeast Region	South Region	Center-West Region
Rural/urban intrastate	24	28	55	81	46
Rural/urban between neighbors	28	30	51	86	47
Rural/urban between non-neighbors	46	47	43	75	58
Rural/rural intrastate	25	26	54	77	41
Rural/rural between neighbors	28	25	49	80	47
Rural/rural between non-neighbors	47	38	45	70	54
Rural non-migrants	23	29	57	85	46

Source: FIBGE, 2000.

Table 4 presents the data for the proportion of married people for migrants and non-migrants for the same categories discussed above. As civil status is highly age dependent, the table below shows data only for the population with age between 20 and 29. Rural/rural flows had greater proportions of married than rural/urban flows and also than non-migrants. Most of the rural/urban flows showed lower values than non-migrants, suggesting that the marriage market of urban centers attract rural dwellers.

Table 4. Comparison between flows of migrants of different types and rural non-migrants for proportion of married for persons with age between 20 and 29 years

Type of flow	Proportion of married for persons with age between 20 and 29 years				
	North Region	Northeast Region	Southeast Region	South Region	Center-West Region
Rural/urban intrastate	18	21	34	38	31
Rural/urban between neighbors	18	23	31	32	27
Rural/urban between non-neighbors	27	26	29	32	29
Rural/rural intrastate	28	29	47	50	37
Rural/rural between neighbors	26	28	39	46	34
Rural/rural between non-neighbors	30	30	40	46	36
Rural non-migrants	21	24	35	38	34

Source: FIBGE, 2000.

The next two tables compare migrants and non-migrants for mean schooling for the population in general and for the group with age between 20 and 29 years. Notice that educational levels are very low in most regions in rural Brazil. The second point to be emphasized is that non-migrants in the North and Northeast regions have schooling levels that are much lower than in the other regions, with slightly smaller values in the Northeast, as can be seen in the last line of the two tables. The South Region had the highest numbers, followed by the Southeast and Center-West. Thirdly, rural/rural migration presents lower levels than rural/urban in all regions. Urban centers do have more educational opportunities than rural areas, especially for higher levels of formal education. Some of this differences in table 5 can be caused by an after migration increase in schooling levels. However, this might not happen significantly with data Table 6, because of the age of the migrants and on account of the short mean time span after migration. In the North and Northeast regions, non-migrants had very low levels of formal education, short step rural/rural migrants had similar levels as the first ones, and rural/urban migrants and interstates between non-neighbors rural/rural migration had higher schooling levels. That is, there was a positive selection for these last types of migration, but not for rural/rural short distance ones. It was also observed a trend of increase in schooling with an enhancement in distance. For the Southeast, South and Center-West regions, it was verified that non-migrants had an education level between the two types of migrants, a little higher than rural/rural ones and lower than the rural/urban. No trend with distance was observed for the first region, and a slight was verified for the last two.

Table 5. Comparison between flows of migrants of different types and rural non-migrants for schooling

Type of flow	Mean schooling (years)				
	North Region	Northeast Region	Southeast Region	South Region	Center-West Region
Rural/urban intrastate	3.35	2.98	4.51	5.01	3.96
Rural/urban between neighbors	3.37	2.89	4.43	5.08	4.25
Rural/urban between non-neighbors	4.31	3.91	4.30	5.23	4.43
Rural/rural intrastate	2.34	1.93	3.29	3.87	3.04
Rural/rural between neighbors	2.46	1.83	3.20	3.95	3.39
Rural/rural between non-neighbors	3.06	2.71	3.29	4.03	3.43
Rural non-migrants	2.02	1.97	3.41	3.86	3.15

Source: FIBGE, 2000.

Table 6. Comparison between flows of migrants of different types and rural non-migrants for schooling for persons with age between 20 and 29 years

Type of flow	Mean schooling for persons with age between 20 and 29 (years)				
	North Region	Northeast Region	Southeast Region	South Region	Center-West Region
Rural/urban intrastate	4.97	4.52	6.22	6.97	5.78
Rural/urban between neighbors	5.09	4.11	5.83	6.96	5.92
Rural/urban between non-neighbors	5.95	4.96	5.96	7.22	6.08
Rural/rural intrastate	3.52	2.90	4.47	5.27	4.22
Rural/rural between neighbors	3.54	2.72	4.35	5.37	4.64
Rural/rural between non-neighbors	4.03	3.73	4.65	5.66	4.40
Rural non-migrants	3.71	3.51	5.45	6.05	5.14

Source: FIBGE, 2000.

This section showed some descriptive data about the selectivity of migration. The next one presents a mathematical simulation that discusses this selectivity.

4. Mathematical Simulation for the Selectivity of Migration

As we have seen in the discussion above, migration is a selective process, that is, migrants are not a random sample of the population. The data above showed that the mean characteristics of the flows are context dependent. From this discussion, follow two questions: which type of region attracts skilled and which one attracts non-skilled population? For which kind of flow, there is a positive and for which there is a negative selection?

In this subsection, it is proposed a mathematical simulation that will discuss this selectivity of migration and the possibility of existence of poverty traps, as described above by Kothary (2002). The simulations are based in two equations, which had as initial motivation some aspects of chemical reactions. The first is a long run one and include properties of the origin and of the destiny of the migrant. It can be related to a thermodynamically determined relation of reactants and products in Chemistry (Carey and Sundberg, 1984). This type of equation indicates the feasibility of migration, as was discussed in the human capital model above, where migration will occur if the net benefits of migration are positive. However, this equation does not address the kinetic of change of residence place, that is, the path of reaction, when there is a transition state between life in origin and in the destiny and migrants are not well established in neither places. This generates an unstable situation, which promotes reversibility and turnover, which are addressed by the second equation of the simulation.

The development of the first of these equations is based in the Roy model. As described in Borjas (1987), the local income distribution and dispersion among human capital levels is the key point to answer the above questions. Generally, the return for schooling is larger and the differences between wages for non-skilled and skilled workers are greater in poor countries. Consequently, the wages for the low skilled in developing countries are normally smaller than the observed in the USA, what may not be true for the workers with higher levels of human capital. If individuals migrate to places with higher salaries, there would be a tendency for the non-skilled to migrate from the poor countries to the USA and this would not happen to higher skilled ones. There would be a negative selection. The contrary would occur for migrants that have as origin a developed country with smaller returns to education and smaller wage dispersion than in the USA. For this type of country, skilled would earn less in the origin than in the USA, but this would not happen for the non-skilled. Therefore, migrants would present a tendency to be positively selected, with mean human capital levels superior than the non-migrants in the origin (Chiquiar and Hanson, 2002).

The above proposition can be described as follows. Wages in each region depend only on the human capital level of the individual and on specific regional parameters. The general equation is described below for three regions:

(2) $\ln(W_i) = \mu_i + \upsilon_i S,$

where i are regions, W_i are wages, $\mu_i > 0$ are the exponential of the salaries for non-skilled workers, $\upsilon_i > 0$ represent the returns for human capital and $S \geq 0$ is the level of the human capital of the person.

In order to understand the implications of the Roy model for the selectivity of migration, suppose that there are three regions with $\mu_3 > \mu_2 > \mu_1$ and $\upsilon_1 > \upsilon_2 > \upsilon_3$. Assume also that migrants will migrate to the region that offers the largest salaries between the origin and region 2, that is, there are no flows between regions 1 and 3. Individuals with low levels of human capital will migrate from region 1 to region 2, because $\mu_2 > \mu_1$. The persons with low levels of schooling in region 3 will prefer not to migrate, because the wages are greater in region 3: $\mu_3 > \mu_2$. For the population with high levels of human capital, the migration would occur between the region 3 and region 2, because the returns to education are greater in region 2. The high skilled in region 1 would not migrate to region 2 due to the higher returns to education in the first one. Consequently, there would be a negative selection in region 1 for the flows with region 2 as destiny, and a positive selection for region 3.

The model presented above can initially be used to illustrate the selectivity of migration, but did not include many aspects that were briefly discussed in the human capital model for migration in the previous section, such as: temporal analyses, costs of migration, discount rates, etc. Aspects of both models are discussed below jointly. First, equation (1) is modified to include wages instead of utilities:

$$\textbf{(3)}\ G_{ij} = \int_0^t (W_{jt} - W_{it})e^{-pt}dt - C_{ij}$$

Equation **(2)** is also rewritten as: **(4)** $W_i = e^{\mu i + \upsilon i S}$.

For simplicity in the mathematical simulation, a polynomial function will be introduced in equation (3) instead of the exponential one in (4), but with the same characteristics, such as $W \geq 0$, $dW/dS \geq 0$, $d^2W/dS^2 \geq 0$ and $d^3W/dS^3 \geq 0$. Equation (4) is approximately rewritten as:

$$W_i = e^{\mu i + \upsilon i S} \approx \alpha_i + \beta_i S + \delta_i S^2 + \sigma_i S^3,$$
where $\alpha_i \geq 0$, $\beta_i \geq 0$, $\delta_i \geq 0$ and $\sigma_i \geq 0$.

The variable S indicates the amount of human capital of the individual. Initially, assume that this amount is a function of time and of spatial parameters. In rural areas, schooling opportunities are smaller than in urban centers and the labor market is less diversified. Hence, human capital levels present a tendency to be smaller in these former ones than in the latter. As the possibility of acquiring higher levels of formal education exists basically in urban centers and also that the labor market of these centers is more sophisticated and demand more human capital, assume that in rural areas the level of human capital is constant and the same the individual had when it migrated: **(5)** $S_r(t) = S_0$. In urban centers, the human capital increases after migration and this increase is constant with time: **(6)** $S_u(t) = S_0 + at$, where $a \geq 0$.

The first basic equation of the simulations is composed of equations (3) to (6). They indicate if the net returns of migration are positive. However, migration will only occur if the returns are positive, but also if the individual can pay the costs of migration in the short run, that is, can overcome the difficulties posed by the transition state, otherwise the person do not have the option to migrate, even if in the long run the returns are positive. Assume that the migrant can migrate only if the differences between earned income and basic costs in a short period of time t´ after migration are larger than the costs of migration. Equation (9) below shows this relation:

$$\textbf{(7)}\ H = \int_0^{t'} [(W_{ij}(S) - E(S)]dt - C_{ij}\ ,$$

where E(s) are the basic costs that can or will not be used to pay migration costs

Generally, lower income individuals have a marginal propensity to consume that is superior to higher income ones. Consequently, the basic costs increase with human capital, but at a lower rate than the increase in income. In the simulation, a polynomial with these characteristics was introduced in (7):

$$E(S) = \phi_d + \varphi_d S_d + \theta_d S_d{}^2\ ,$$
where $0 \leq \phi_d \leq \alpha_d, 0 \leq \varphi_d \leq \beta_d, 0 \leq \theta_d \leq \delta_d$,

Equations (3) to (6) form the long-term equation, while equation (7) represents the short term one, both include the costs of migration. This costs are a function of the distance, d_{ij} , of the type of migration, k_i and of the effectiveness of the social nets of the potential migrant between two specific localities, R_{ij}, that is, $C_{ij}^{k} = C_{ij}^{k}(d_{ij}, k_i, R_{ij})$. This function will be written as: **(8)** $C_{ij}^{k} = D(d_{ij})K(k_i)(1-R(r_{ij}))$.

As was discussed above, the costs of migration increase with the distance. A common used equation in aggregated studies for this relation is given by: $D_{ij}(d_{ij}) = Ad_{ij}^{\alpha}$, where A and α are positive constants and $\alpha \in (0,1)$. As we discuss three types of distance in the empirical analyses, in these simulations, $D(d_{ij})$ has only three values, A, 3/2A and 4A respectively for intrastate, interstate between neighbors and interstate between non-neighbors migration. These values are an approximation for the mean value of the function of the distance between the origin and the destiny of the migrant for each one of these types of migration.

If human capital that was acquired by the migrant in the origin can be used effectively in a short time span in the destiny, the costs of migration might be smaller than otherwise. Consequently, it can be said that migrations between localities with similar labor markets may present lower costs than other types, and, therefore, rural/rural migrations may have lower costs than rural/urban ones, if all other variables are held constant. The $K(k_i)$ function deals with these differences in costs that are not related to distance or social nets, and will assume the arbitrary values of 3/2 and 1 respectively for the rural/urban and rural/rural types of migration.

As cited above, the existence of an effective social net may diminish decisively the costs of migration. It is reasonable to think that the effectiveness of a social net is a function of the number of individuals in the migrants' potential destiny that had as origin the same place of the migrants´ present locality of residence. This number is represented by r_{ij}. $R = R(r_{ij})$ is a function with the following characteristics: $R(0) = 0$, $R'(r) > 0$, $R''(r) < 0$ and $\lim_{r\to\infty} R = 1$, as 1 - $R(r_{ij})$ is the effectiveness of cost decrease due to the social net. A function with these characteristics is **(9)** $R_{ij}(r_{ij}) = 1 - e^{-wr_{ij}}$, where r_{ij} as cited, is the proportion of the population of **j** that had **i** as origin, **w** is the effectiveness of the links between migrants and non-migrants in the origin.

All these equations are based on empirical findings. In order to illustrate their consequence in the selectivity of migration some simulations are presented below, first for the rural/rural migration and then for rural/urban ones.

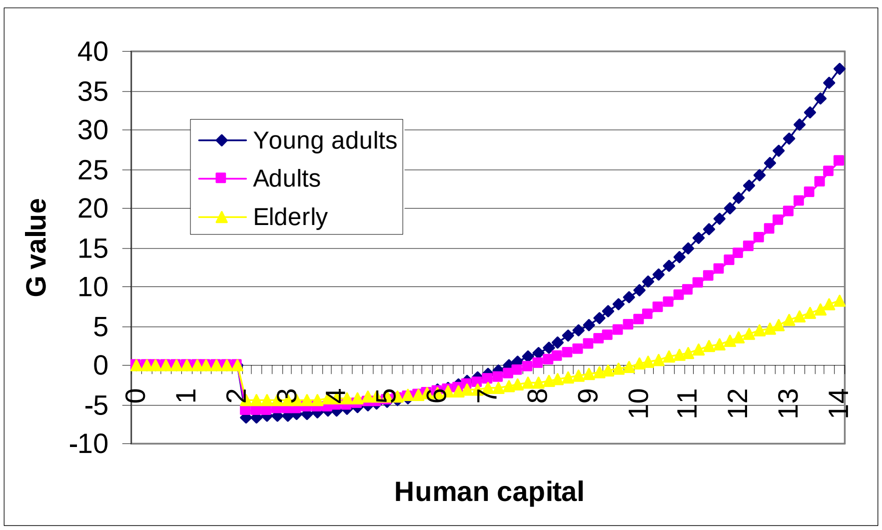
4A. Rural/rural migration simulations

The above equations were described for rural/rural and rural/urban migration. In particular for this first type of migration, as was described above, $S_d(t) = S_o(t) = S_0$, that is, the human capital does not increase after migration. As migration is a rural/rural one: $K(k_i)$ = 1 and, consequently, $C_{ij}^{r/r} = AD(d_{ij})(e^{-wr})$. The final equations for G and H are obtained by introducing these particularities in equations (3) and (7) and integrating.

As discussed above, migration will occur if both, G and H, are positive. A initial simulation was done with arbitrary values that were chosen to be a benchmark for comparisons. For the G equation, the chosen values were the following. The discount rate is $\rho = 0.02$. Time, t, is defined as t = 70 – age, that is approximately the time spent in the destiny after migration, assuming a mean life duration of 70 years. The wage coefficients for the destiny are similar to the ones obtained for urban Brazil in 2004: $\alpha_d = 1.4$, $\beta_d = 0.1$, $\delta_d = 0.0002$ and $\sigma_d = 0.0005$. For the origin, the values are similar to rural Brazil in this same year: $\alpha_o = 1.5$, $\beta_o = 0.1$, $\delta_o = 0.0$ and $\sigma_o = 0.0$. The constants in the costs function were chosen to obtain H > 0 and G > 0 for reasonable values of human capital. These are: A = 10; as already mentioned, D = A, for intrastate migrations, D = 3A/2 for interstate between neighbors, and D = 4A for interstate between non-neighbors; $\omega = 1$; and r = 1. Regarding the short run equation, the chosen values for the constants are: $\phi_d = 1.2$, $\varphi_d = 0.05$, $\theta_d = 0$ and $t' = 12$.

Initially, as presented in diagram 1 with G values, three simulations were done with these values for young adults, age = 20, adults, age = 40, and elderly, age = 60 for intrastate migration. As can be seen in the diagram by the G = 0 value, migrants can only pay the costs of migration in the short run if their human capital is $S \geq 2.2$. Below this value they cannot afford to migrate independently of the G value and are trapped in their origin. If $S \geq 2.2$, as H > 0, the individual has the option whether or not to migrate. Tracking the long run equation, migration will occur if G is also above 0. As can be seen in the diagram, for young adults, this happens if $S \geq 7.6$. The same is verified for adults, if $S \geq 8.0$, and for the elderly, if $S \geq 10.0$. This simulation is an example of positive selection. Notice that $\alpha_o > \alpha_d$, non-skilled earn more in the origin and even if they could afford, they would not migrate.

Diagram 1. Net returns of migration



The results obtained in the above simulation depend directly on the chosen values for the constants and should be appreciated as a point of reference. Other simulations are shown in tables below so that the implications can be better understood. The first one compares intrastate with interstate between neighbors and between non-neighbors migrations. The constants have the same value as presented above, but the K value varies, respectively K = A, K = 3A/2 and K = 4A. With the increase of the distance, the costs of migration also raise and the same takes place with the needed S values. Observe that for the intrastate migration, the short-term equation will be positive for S ≥ 2.2. For the interstate between neighbors migration, the same will occur for S ≥ 4.4 and for the non-neighbors migration for a much higher value, S = 10.2. Following this simulation, individuals with very low human capital, below S = 2.2, cannot afford to migrate at all and will be non-migrants. Persons with a relative low human capital level, between S = 2.2 and S = 10.2, can afford to migrate in the same or to a neighbor state, but cannot pay the high costs of distant migration. However, as is shown by the threshold of the G function, individuals with human capital level between S = 2.2 and S = 7.6 can afford to migrate, but the net returns of migration is not positive, due to the higher wage levels in the origin for the non-skilled. Only for individuals with a medium level of human capital, S ≥ 7.6, roughly a complete degree of fundamental schooling, migration will occur. Only persons that hold approximately a High School degree (S ≥ 10.4) can migrate to a distant locality. Notice that for the elderly, the migration between non-neighbors is not possible, because G is negative for any value of human capital.

Table 7. Threshold values for H and G functions – rural/rural migration

Simulations	H > 0	G > 0		
		Young adults	Adults	Elderly
Intrastate	≥ 2.2	≥ 7.6	≥ 8.0	≥ 10.0
Interstate between neighbors	≥ 4.4	≥ 8.2	≥ 8.8	≥ 11.2
Interstate between non-neighbors	≥ 10.2	≥ 10.4	≥ 11.4	-

If the social net is more effective in diminishing the costs of migration or if the proportion of the population in the destiny with the same origin of the individual is increased, which would be the consequences? For these simulations, the value for the product $\mathbf{wr_{ij}}$ was multiplied by four. Observe that the threshold values are much lower, especially for H. If the social net is as effective as in this simulations, even persons with extremely low human capital levels can afford to migrate, at least in short distances movements, although they will not migrate due to the negative values for the net returns. Notice that the elderly would migrate in this situation, even for distant values. This may explain part of what happens with the return migration after retirement.

Table 8. Threshold values for H and G functions – changes in ω and r values

Simulations	H > 0	G > 0		
		Young adults	Adults	Elderly
Intrastate	≥ 0.0	≥ 6.6	≥ 6.8	≥ 7.8
Interstate between neighbors	≥ 0.0	≥ 6.8	≥ 7.2	≥ 8.6
Interstate between non-neighbors	≥ 4.4	≥ 8.2	≥ 8.8	≥ 11.2

In the next simulations, the coefficient for wages of non-skilled workers in the destiny increases from α_d = 1.4 to α_d = 1.55. All other variables values did not change. Notice that this small change enables the unskilled to migrate in an intrastate migration.

Table 9. Threshold values for H and G functions – changes in α_d values

Simulations	H > 0	G > 0		
		Young adults	Adults	Elderly
Intrastate	≥ 0.0	≥ 0.0	≥ 2.6	≥ 5.8
Interstate between neighbors	≥ 0.0	≥ 3.0	≥ 4.2	≥ 7.0
Interstate between non-neighbors	≥ 2.0	≥ 6.2	≥ 7.2	≥ 10.2

There are many possibilities of simulations. Here some of them were presented as illustration. Some others are shown for the rural/urban migration.

4B. Rural/urban migration simulations

Now we turn our attention to the rural/urban migration. Some modifications must be done in the simulations. The human capital in the destiny is now a function of time: $S_o(t) = S_0$ and $S_d(t) = S_0 + at$. The costs of migration also change because origin and destiny are no longer similar: $K(k_i) = 3/2$.

The values are the same as the first one of the simulations: $\rho = 0.02$, $\alpha_d = 1.4$, $\beta_d = 0.1$, $\delta_d = 0.0002$, $\sigma_d = 0.0005$, $\alpha_o = 1.5$, $\beta_o = 0.1$, $\delta_o = 0.0$, $\sigma_o = 0.0$, $A = 10$, $\omega = 1$, $r = 1$, $\phi_d = 1.2$, $\varphi_d = 0.05$ and $\theta_d = 0$. The initial value for **a** is zero. So, the only difference between this simulation and the first one presented is that K is no longer one, as in the rural/rural migration. Observe that the H and G threshold values are larger in table 10 than in table 7, especially for the short-term equation in the short distance migrations. This indicates that very low skilled migrants may be able to migrate from rural areas to similar ones, but not to urban centers.

Table 10. Threshold values for H and G functions – rural/urban migration

Simulations	H > 0	G > 0 Young adults	Adults	Elderly
Intrastate	≥ 4.4	≥ 8.2	≥ 8.8	≥ 11.2
Interstate between neighbors	≥ 6.8	≥ 9.0	≥ 9.8	≥ 12.6
Interstate between non-neighbors	≥ 12.6	≥ 12.6	≥ 12.8	-

What would happen if the migrant could acquire extra human capital in urban centers? In order to test this, **a** is increased to a = 0.1. That means that for each year in the urban center, the migrant increase 0.1 units (roughly schooling years) of human capital. As can be seen by the results in table 11, this enables lower skilled young adults to migrate, particularly for the short migrations, when compared with the ones in table 10.

Table 11. Threshold values for H and G functions – changes in a values

Simulations	H > 0	G > 0 Young adults	Adults	Elderly
Intrastate	≥ 3.0	≥ 3.0	≥ 6.0	≥ 10.4
Interstate between neighbors	≥ 5.2	≥ 5.2	≥ 7.2	≥ 12.2
Interstate between non-neighbors	≥ 11.2	≥ 11.2	≥ 11.2	-

In the next simulation, the saving power of the migrant is increased. The new values for two of the constants are: $\phi_d = 1.0$ and $\varphi_d = 0.0$. Besides this, there is a small increase in the rate of growth of human capital, a = 0.12. Table 12 presents the results. It can be seen that these changes turns feasible for the very low skilled to migrate in an intrastate migration for young adults.

Table 12. Threshold values for H and G functions – changes in saving power

Simulations	H > 0	G > 0 Young adults	Adults	Elderly
Intrastate	≥ 0.2	≥ 0.8	≥ 5.4	≥ 10.2
Interstate between neighbors	≥ 1.8	≥ 3.2	≥ 6.8	≥ 11.8
Interstate between non-neighbors	≥ 7.2	≥ 7.6	≥ 10.6	-

In all the above simulations, there were positive selections. When can it be a negative or an intermediate one? For instance, a simulation was done with the main difference that low skilled wages are higher in the destiny than in the origin: $\alpha_d = 2.0$ and $\alpha_o = 1.5$. For the intrastate migration, due to the low costs of migration, these differences promoted the migration of them, but not for the higher levels of human capital individuals, in a negative selection of migrants.

With the same constants, it would be observed the intermediate selection for the interstate migration between neighbors. This happens because the very low skilled can not pay the costs of migration, and for the skilled, the net returns of migration are negative.

These simulations enlighten some points about the selectivity of migration in Brazil. As was seen in the previous section, in the North and Northeast regions where the non-migrants have very low levels of schooling, and consequently of income, a large proportion of the population may be below the threshold of migration, especially for the long distance and rural/urban ones. Some points of these simulations are empirically analyzed and the results are showed in the following sections. The next one presets the methodology and data, and the following discusses the empirical results.

5. Methodology and Data

In this paper, the selectivity of the migratory process was empirically analyzed with the application of multinomial logistic micro models. The objective is to identify personal attributes that modify the individual probabilities of being a non-migrant of a migrant of different types.

In the multinomial logistic model, the probability that a specific event **j** will occur for individual **i** is given by:

$$\textbf{(10)}\quad \Pr ob(Y_i = j) = \frac{e^{\beta_j' X_i}}{1 + \sum_h e^{\beta_h' X_i}},$$

where X_i is the vector of individual characteristics, β_is are the regression coefficients that were obtained for all the analyzed events **h**, including **j**, a particular one, and the coefficients β_k of the event being used as comparison standard, here to be a non-migrant, are all equal to zero.

In this model, the logarithm of the odds ratio can be estimated by the following equation:

$$\textbf{(11)}\quad \ln\left(\frac{P_{ij}}{P_{ik}}\right) = X_i'(\beta_j - \beta_k) = X_i'\beta_j,$$

where P_{ij} is the probability that event **j** will occur for the individual **i** (in this particular study is the probability that the person will migrate in one of the types of migration mentioned above) and P_{ik} is the probability that event **k** will happen for the same individual (here is to be a non-migrant). One basic assumption is that the probability of one possibility will not impact on the others probabilities.

The micro data of the Brazilian Demographic Census of 2000 was used as database. The migrants with rural origin and rural non-migrants were selected. Only the individuals with age between 18 and 64 were included in the analyses, most of the individuals that migrate because of labor market characteristics. The multiple regressions were done separately for each region, always comparing the migrants with the non-migrants in the same origin.
The response variable has 7 categories. The individual can be a non-migrant or a migrant. If the person is a migrant it can be intrastate, interstate between neighbors or interstate between non-neighbors migrant of one of two types – rural/rural or rural/urban. The dependent variable had

always the non-migrant category as a standard for comparison. The independent variables were: age (years), age squared, sex (1 for male and 0 for female), ethnic group (1 for White/Asiatic and 0 for Black/Pardo/Indigenous), civil status (1 for married and 0 otherwise) and schooling (years of formal education).

The main empirical results are shown below.

6. Empirical Results

The empirical results are presented in two subsections. The first shows the results that were obtained for the multinomial logistic models in ten tables. The other, based on these regressions, presents some simulations that illustrate the obtained results.

6.A. Multinomial logistic models

In this subsection, the results are presented in two tables for each region: the first one with the rural/urban flows and the second with the rural/rural ones. The coefficient value (Coef.), standard deviation (S.D.) and odds ratio (odds) are shown. Notice that the great majority of the coefficients were significant, a few of them, that are presented bold face, were not.

Table 13 presents the results for the rural/urban intrastate migration in the North Region are in the second, third and fourth columns. The age coefficient showed a negative sigh and the coefficient for age squared was positive, meaning that the probability of being a migrant decreased with age for young adults, as was observed in the mathematical simulation, and in a convex manner. De to this convexity, the probability increased after a minimum. Life cycle aspects, such as retirement, can explain this increase. Notice that in all models for all macroregions the same result was obtained and no further commentary is given for these variables.

The sex dummy coefficient showed a negative sign, indicating that being a man decreased the probability of being a migrant. As can be seen for this variable, the odds ratio was 0.83, much smaller than one. Women are normally relatively more attracted to urban destinies, what partly explains this result. The coefficient for the ethnic group was not significant, that is, to be a White/Asiatic did not impact on the probability of being a rural/rural intrastate migrant. For civil status, the coefficient was negative, indicating that to be married decreased the probability of being a rural/urban migrant. The observable fact of short migration to urban areas in order to get married does exist in some regions.

Notice that, when the intrastate migration is compared to the interstate between neighbors and to between non-neighbors migrations, some differences are observed for these three last variables. The coefficients for the sex, the ethnic group and the civil status dummy were negative or not significant for short distance migrations, but positive for the interstate between non-neighbors one with a clear tendency of increase with distance, indicating that, after the inclusion of the other variables in the model, to be a men, White/Asiatic and married increased the probability of being a migrant for longer steps.

In the mathematical simulations, especially due to the short run costs of migration, it was proposed that higher income groups might be relatively more capable of migrating in a long distance step. This fact was empirically analyzed by the schooling variable. Notice that all the coefficients were positive, especially for the interstate between non-neighbors migration, as expected, indicating the higher levels of schooling of these migrants when compared to rural non-migrants.

Table 13. Multinomial logistic model – rural/urban flows – North Region

Variable	Intrastate			Interstate between neighbors			Interstate between non-neighbors		
	Coef.	S.D.	Odds	Coef.	S.D.	Odds	Coef.	S.D.	Odds
Intercept	3.439	0.0146	-	2.138	0.0213	-	0.826	0.0293	-
Age	-0.363	0.0010	-	-0.363	0.0014	-	-0.367	0.0019	-
Age squared	0.004	0.0000	-	0.004	0.0000	-	0.004	0.0000	-
Sex	-0.190	0.0060	0.83	-0.106	0.0105	0.90	0.033	0.0140	1.03
Ethic group	**-0.006**	**0.0071**	**0.99**	0.190	0.0118	1.21	0.910	0.0141	2.49
Civil status	-0.038	0.0082	0.96	**-0.021**	**0.0150**	**0.98**	0.357	0.0184	1.43
Schooling	0.070	0.0010	1.07	0.067	0.0018	1.07	0.148	0.0022	1.16

Note 1: The results in bold face are not significant at 5%.
Note 2: Non-migrant category was the base for comparison.
Note 3: -2log(likelihood): with intercept only 828520; and final model 468886. Note 4: number of observations = 1944790.

Next table shows the results for the rural/rural migration. The sex dummy was positive for the two extreme distances, indicating the predominance of males in the rural/rural migration, contrary to the observed for rural/urban short distance migrations. The ethnic coefficient showed the same tendency of the table above: smaller coefficients for short steps and larger for longer. For the civil status, all the coefficients were positive, indicating that to be a married person increased the probability of being a rural/rural migrant, what was observed for rural/urban migration only for long distances. This suggests that to some degree the urban marriage market attract single rural dwellers, but not the married ones. The schooling coefficient, contrary to the observed above for the rural/urban migration, showed a negative sigh for short distances, but a positive for longer. This suggest that the threshold for migration, as explained in the mathematical simulations, are not very effective for rural/rural short migrations, that is, these types of migration are accessible for most of the rural population, as expected owing to the lower cost of similar short migrations.

Table 14. Multinomial logistic model – rural/rural flows – North Region

Variable	Intrastate			Interstate between neighbors			Interstate between non-neighbors		
	Coef.	S.D.	Odds	Coef.	S.D.	Odds	Coef.	S.D.	Odds
Intercept	2.994	0.0155		1.278	0.0274		-0.310	0.0493	
Age	-0.326	0.0010		-0.318	0.0019		-0.342	0.0033	
Age squared	0.004	0.0000		0.004	0.0000		0.004	0.0000	
Sex	0.089	0.0067	1.09	**0.018**	**0.0137**	**1.02**	0.089	0.0248	1.09
Ethic group	0.103	0.0077	1.11	0.285	0.0151	1.33	1.013	0.0248	2.75
Civil status	0.415	0.0086	1.51	0.378	0.0182	1.46	0.580	0.0319	1.79
Schooling	-0.086	0.0013	0.92	-0.073	0.0028	0.93	0.012	0.0045	1.01

Notes: Same as above.

Similar tables as the two above are shown for the other four macroregions in Brazil. Table 15 presents the results for the rural/urban flows for the Northeast Region. Firstly, notice the similarity with table 13, for the same flows in the North Region. The coefficients for age, age squared, sex and schooling were similar. Emphasizing these last ones, as observed for the North Region, all the coefficients were positive, indicating that higher levels of schooling increased the probability of being a migrant. When the results of these two macroregions are compared, the major differences, although small, were observed for three coefficients: for ethnic group for short distances migrations and for civil status in the interstate between neighbors. The same general tendency was observed, for both coefficients, they increased with the distance, but it was negative for intrastate migration and non-significant for the between neighbors one for ethnic group and positive for the civil status for the between neighbors migration.

Table 15. Multinomial logistic model – rural/urban flows – Northeast Region

Variable	Intrastate			Interstate between neighbors			Interstate between non-neighbors		
	Coef.	S.D.	Odds	Coef.	S.D.	Odds	Coef.	S.D.	Odds
Intercept	3.416	0.0081		1.776	0.0139	0.00	1.539	0.0122	0.00
Age	-0.357	0.0005	0.70	-0.352	0.0009	0.70	-0.328	0.0008	0.72
Age squared	0.004	0.0000	1.00	0.004	0.0000	1.00	0.004	0.0000	1.00
Sex	-0.264	0.0033	0.77	-0.165	0.0069	0.85	0.022	0.0050	1.02
Ethic group	-0.071	0.0036	0.93	**0.012**	**0.0075**	**1.01**	0.726	0.0050	2.07
Civil status	-0.031	0.0043	0.97	0.104	0.0093	1.11	0.304	0.0066	1.36
Schooling	0.029	0.0006	1.03	0.008	0.0012	1.01	0.088	0.0008	1.09

Note 1: The results in bold face are not significant at 5%.
Note 2: Non-migrant category was the base for comparison.
Note 3: -2log(likelihood): with intercept only 2383356; and final model 1063728. Note 4: number of observations = 7918930.

Notice that the rural/urban and the rural/rural flows that originated in the Northeast (tables 15 and 16) presented many similarities. These include the coefficients for the sex dummy, which were negative for short distance flows, indicating the high proportion of women in the flows of the macroregion, as was verified in table 2 and positive for non-neighbors migrations. The coefficients for ethnic group also presented the same general trend of increase with distance. Besides these resemblances, there are two main differences between these types of flows. The first one is the positive coefficient for civil status for the rural/rural intrastate migration, contrary to the observed for rural/urban, indicating that being a married person increased the probability of being a short distance migrant of this first type, as was verified for the North Region. And, most importantly, all the negative coefficients for schooling for all distances, a rather different situation than observed for the rural/urban flows, indicating that to have lower levels of formal education increased the probability of being a rural/rural migrant.

Table 16. Multinomial logistic model – rural/rural flows – Northeast Region

Variable	Intrastate			Interstate between neighbors			Interstate between non-neighbors		
	Coef.	S.D.	Odds	Coef.	S.D.	Odds	Coef.	S.D.	Odds
Intercept	3.250	0.0088	0.00	1.504	0.0167	0.00	0.371	0.0232	0.00
Age	-0.341	0.0006	0.71	-0.337	0.0011	0.71	-0.333	0.0016	0.72
Age squared	0.004	0.0000	1.00	0.004	0.0000	1.00	0.004	0.0000	1.00
Sex	-0.128	0.0038	0.88	-0.066	0.0088	0.94	0.113	0.0113	1.12
Ethic group	-0.126	0.0043	0.88	-0.151	0.0101	0.86	0.403	0.0115	1.50
Civil status	0.215	0.0050	1.24	0.303	0.0118	1.35	0.492	0.0147	1.64
Schooling	-0.134	0.0008	0.87	-0.160	0.0020	0.85	-0.029	0.0021	0.97

Notes: Same as above.

The next 6 tables present the results for the other macroregions in Brazil and will be discussed jointly with the four above. As already pointed out, all the coefficients for age and age squared were similar in all models. The coefficients for the sex dummy indicated some general trends. The rural/urban flows tend to show smaller values, that is, they are more feminine, as expected, and there is a tendency for the coefficients to increase with distance, most turning positive for long distance steps, suggesting the distinct propensity to migration between the sexes for different distances. The majority of the coefficients for rural/rural migration are positive, indicating the greater inclination for male migration between rural areas. However, for short distance steps in the Northeast this did not occur.

Table 17. Multinomial logistic model – rural/urban flows – Southeast Region

Variable	Intrastate			Interstate between neighbors			Interstate between non-neighbors		
	Coef.	S.D.	Odds	Coef.	S.D.	Odds	Coef.	S.D.	Odds
Intercept	3.233	0.0100		1.888	0.0159	0.00	0.631	0.0291	0.00
Age	-0.308	0.0006	0.74	-0.298	0.0010	0.74	-0.303	0.0019	0.74
Age squared	0.004	0.0000	1.00	0.003	0.0000	1.00	0.004	0.0000	1.00
Sex	-0.149	0.0036	0.86	-0.146	0.0071	0.86	-0.114	0.0145	0.89
Ethic group	-0.111	0.0037	0.90	-0.212	0.0072	0.81	-0.538	0.0147	0.58
Civil status	-0.027	0.0044	0.97	-0.136	0.0091	0.87	-0.269	0.0189	0.76
Schooling	**0.016**	**0.0006**	**1.02**	**-0.004**	**0.0011**	**1.00**	**-0.003**	**0.0023**	**1.00**

Note 1: The results in bold face are not significant at 5%.
Note 2: Non-migrant category was the base for comparison.
Note 3: -2log(likelihood): with intercept only 1320302; and final model 675892. Note 4: number of observations = 4197819.

Table 3 showed that the proportions of White/Asiatic are very heterogeneous between the macroregions in Brazil. This is expected to echo in the coefficients of the multinomial regressions, especially the between non-neighbors flows. The three regions with lower proportions of White/Asiatic, North, Northeast and Center-West, showed the same tendency of increase in the coefficient of ethnic group, indicating that the propensity to be a long distance migrants increase for White/Asiatic. This positive coefficient for long distance migrants can be explained by at least two facts. The first one is that due to the higher cost of migration, individuals of this group can cope more effectively with them because they have higher mean income than persons of the Black/Pardo/Indigenous group. Another point is that some of the long distance migrants with origin in one of these three regions maybe return migrants to the South and Southeast regions, where the population present grater proportions of White/Asiatic. Conversely, for these two last region, it can be seen that the coefficients are mostly negative and decreasing, suggesting the return and high turnover rate of migrants with origin in other regions. Just one type of migration in the South Region, the rural/urban between neighbors, because of the interchange with other countries and São Paulo state, all with White/Asiatic majorities that are numerous, differed from the others.

Table 18. Multinomial logistic model – rural/rural flows – Southeast Region

Variable	Intrastate			Interstate between neighbors			Interstate between non-neighbors		
	Coef.	S.D.	Odds	Coef.	S.D.	Odds	Coef.	S.D.	Odds
Intercept	2.945	0.0115	0.00	1.230	0.0224	0.00	0.348	0.0333	0.00
Age	-0.293	0.0007	0.75	-0.288	0.0015	0.75	-0.291	0.0023	0.75
Age squared	0.003	0.0000	1.00	0.003	0.0000	1.00	0.003	0.0000	1.00
Sex	0.052	0.0048	1.05	0.077	0.0113	1.08	0.156	0.0171	1.17
Ethic group	**-0.007**	**0.0048**	**0.99**	-0.194	0.0113	0.82	-0.400	0.0171	0.67
Civil status	0.340	0.0059	1.40	0.103	0.0145	1.11	0.181	0.0216	1.20
Schooling	-0.127	0.0009	0.88	-0.137	0.0021	0.87	-0.116	0.0031	0.89

Notes: Same as above.

The marriage market of the urban centers seams very attractive for migrants, as is suggested by the most important tendencies observed that were the negative sighs for rural/urban migration and the positive ones for the rural/rural. For the North and Northeast, it was observed a tendency of increase in the coefficients with distance, as already discussed, that was not observed in the other regions. This indicates that couples may present lower relative cost of migration for longer steps with origin in one of those regions, the ones with the lowest levels of schooling and income.

Table 19. Multinomial logistic model – rural/urban flows – South Region

Variable	Intrastate			Interstate between neighbors			Interstate between non-neighbors		
	Coef.	S.D.	Odds	Coef.	S.D.	Odds	Coef.	S.D.	Odds
Intercept	3.631	0.0125		1.529	0.0254	0.00	1.987	0.0227	0.00
Age	-0.305	0.0007	0.74	-0.309	0.0015	0.73	-0.304	0.0014	0.74
Age squared	0.003	0.0000	1.00	0.003	0.0000	1.00	0.003	0.0000	1.00
Sex	-0.124	0.0042	0.88	-0.056	0.0106	0.95	0.020	0.0096	1.02
Ethic group	-0.245	0.0055	0.78	0.095	0.0151	1.10	-0.646	0.0112	0.52
Civil status	**-0.002**	**0.0052**	**1.00**	**-0.004**	**0.0139**	**1.00**	0.078	0.0123	1.08
Schooling	0.009	0.0007	1.01	**0.000**	**0.0017**	**1.00**	0.028	0.0015	1.03

Note 1: The results in bold face are not significant at 5%.
Note 2: Non-migrant category was the base for comparison.
Note 3: -2log(likelihood): with intercept only 1098420; and final model 552794. Note 4: number of observations = 3104775.

Two general trends were observed for the schooling coefficients. Firstly, most of the rural/urban ones are positive, indicting a positive selection of migrants. Notice that the sample contains individuals with age between 18 and 64 and the mean time in the destiny after migration is around two years, a small period to change significantly schooling levels for this age group. On the other hand, for rural/rural migration, the coefficients were mostly negative, suggesting the negative selection, at least for the observable attributes. For the macroregions with lower levels of schooling, North, Northeast and Center-West, the tendency of increase in the coefficients was observed, what was not verified for the other two regions. These three trends suggest that individuals that migrate to urban areas are very differently than the ones that do so to rural regions. Besides that, and most importantly, individuals with very low educational levels, and consequently of income, do migrate, they are not trapped in their origins, but mostly in short steps rural/rural migrations. As was discussed in the mathematical simulations, they can afford to migrate to close similar area with low cost associated to a change of place of residence.

Table 20. Multinomial logistic model – rural/rural flows – South Region

Variable	Intrastate			Interstate between neighbors			Interstate between non-neighbors		
	Coef.	S.D.	Odds	Coef.	S.D.	Odds	Coef.	S.D.	Odds
Intercept	3.709	0.0137	0.00	1.231	0.0342	0.00	1.777	0.0280	0.00
Age	-0.291	0.0008	0.75	-0.299	0.0022	0.74	-0.286	0.0019	0.75
Age squared	0.003	0.0000	1.00	0.003	0.0000	1.00	0.003	0.0000	1.00
Sex	**-0.004**	**0.0051**	**1.00**	**0.022**	**0.0160**	**1.02**	0.196	0.0130	1.22
Ethic group	-0.400	0.0062	0.67	-0.233	0.0200	0.79	-0.787	0.0141	0.46
Civil status	0.388	0.0065	1.47	0.279	0.0210	1.32	0.373	0.0167	1.45
Schooling	-0.122	0.0009	0.89	-0.111	0.0029	0.90	-0.100	0.0023	0.90

Notes: Same as above.

Table 21. Multinomial logistic model – rural/urban flows – Center-West Region

Variable	Intrastate			Interstate between neighbors			Interstate between non-neighbors		
	Coef.	S.D.	Odds	Coef.	S.D.	Odds	Coef.	S.D.	Odds
Intercept	2,965	0,0183		1,804	0,0263	0,00	0,894	0,0337	0,00
Age	-0,277	0,0011	0,76	-0,284	0,0017	0,75	-0,281	0,0022	0,75
Age squared	0,003	0,0000	1,00	0,003	0,0000	1,00	0,003	0,0000	1,00
Sex	-0,113	0,0072	0,89	-0,140	0,0124	0,87	**0,027**	**0,0162**	**1,03**
Ethic group	0,032	0,0073	1,03	0,074	0,0125	1,08	0,504	0,0164	1,65
Civil status	-0,094	0,0086	0,91	-0,169	0,0156	0,84	-0,045	0,0202	0,96
Schooling	**-0,001**	**0,0012**	**1,00**	0,022	0,0019	1,02	0,026	0,0025	1,03

Note 1: The results in bold face are not significant at 5%.
Note 2: Non-migrant category was the base for comparison.
Note 3: -2log(likelihood): with intercept only 524370; and final model 352734. Note 4: number of observations = 957544.

Table 22. Multinomial logistic model – rural/rural flows – Center-West Region

Variable	Intrastate			Interstate between neighbors			Interstate between non-neighbors		
	Coef.	S.D.	Odds	Coef.	S.D.	Odds	Coef.	S.D.	Odds
Intercept	2.740	0.0193	0.00	1.215	0.0319	0.00	0.421	0.0439	0.00
Age	-0.242	0.0012	0.78	-0.259	0.0021	0.77	-0.267	0.0029	0.77
Age squared	0.003	0.0000	1.00	0.003	0.0000	1.00	0.003	0.0000	1.00
Sex	0.114	0.0079	1.12	0.153	0.0157	1.17	0.091	0.0226	1.10
Ethic group	-0.050	0.0079	0.95	0.166	0.0156	1.18	0.415	0.0225	1.51
Civil status	0.139	0.0093	1.15	0.102	0.0194	1.11	0.067	0.0282	1.07
Schooling	-0.125	0.0014	0.88	-0.082	0.0028	0.92	-0.076	0.0040	0.93

Notes: Same as above.

In order to illustrate the implications of the results discussed in these multinomial regressions, some simulations are presented below with the coefficients of the Northeast Region, the one with the lowest levels of formal education.

6.B. Illustrative simulations

Equation (11) was used to estimate the probabilities of occurrence of the 7 possibilities that were analyzed in the above equations: six types of migrants and the non-migrants. The βs coefficients are the ones presented in tables 15 and 16. The vector X_i was arbitrary chosen for 6 specific types of individuals. The details for each of them are showed below table 23. Notice that the question being answered here is: if the migrant has the following characteristics, which are the probabilities for the seven possibilities discussed.

The first simulation was done for a typical migrant in the Northeast Region: a non-White/Asiatic bachelor woman with 24 years old and with two years of formal education. Table 23 shows the results for all the 7 possibilities for this simulation in the second column, as described in the descriptive section of this paper. Notice that most individuals with these characteristics would be non-migrants (85.8%). Among the migrants, 14.2%, less than half would be of the rural/rural type (5.99%), as the majority would have an urban destiny (8.20%), and a large majority would be intrastate (10%).

The second simulation was done for a Non-White/Asiatic married woman with 30 years old with two years of formal education. That is, the difference from the first one is that the woman is older and got married. The differences were quite large from the above simulation, indicating the lower mobility mostly due to the increase in age. The probability of migration decreased to less than 10% (8.2%), approximately half of the migrants would have rural destiny (3.9%), and also as above, most would be an intrastate migrant (5.3%).

A third simulation was done with just one modification that was the sex. The probability of being a migrant decreased slightly, from 8.2 to 7.3%, demonstrating the greater mobility of married women when compared to married men, especially for the short distance rural/urban migration. In the fourth simulation, the age was increased from 30 to 50 and schooling was decreased from 2 to 1. All other variables did not change. This modification in educational level is expected in real world as older generations have lower levels of schooling. As can be see, the probability of being a migrant is very low, around 5%, most of it rural/rural and short distance, suggesting, as discussed in the mathematical simulations, that most individuals from older generation do not afford to migrate in longer steps to urban areas.

The last two simulations are done with individuals with very high schooling levels for rural Northeast in 2000, eight years of formal education. Both simulations were done for married man with 30 years old, the first one for non-White/Asiatic and the second for a White/Asiatic. Comparing simulation 5 with the number 3 one, which differ only in schooling level, the probability of migration decreased slightly, from 7.3% to 6.8%. This happened because the probability of a rural/rural migration decreased, while the rural/urban increased, but a little less. The last simulation is similar to the fifth one with the only change of ethic group. Notice that the probability of being a migrant, especially of long distance rural/urban, typically to São Paulo, increased remarkably. These last simulations introduce a point that was not discussed lengthy. In rural Northeast Region, the schooling levels are very low and, surely, they are even lower for non-White/Asiatic making the fifth simulation relatively rare in the real world.

Table 23. Fictitious simulations

Possibility	Simulations and probabilities (%)					
	1	2	3	4	5	6
Rural/rural intrastate migrant	4.73	2.96	2.63	2.31	1.18	1.02
Rural/rural interstate between neighbors migrant	0.86	0.60	0.57	0.57	0.22	0.18
Rural/rural interstate between non-neighbors migrant	0.40	0.34	0.39	0.33	0.33	0.48
Rural/urban intrastate migrant	5.27	2.35	1.82	0.84	2.18	1.98
Rural/urban interstate between neighbors migrant	1.11	0.58	0.50	0.26	0.52	0.52
Rural/urban interstate between non-neighbors migrant	1.82	1.35	1.39	1.02	2.37	4.79
Migrant	14.2	8.19	7.30	5.32	6.81	8.97
Non-migrant	85.8	91.8	92.7	94.7	93.2	91.0

(1) Non-White/Asiatic bachelor woman, 24 years old, with two years of formal education
(2) Non-White/Asiatic married woman, 30 years old, with two years of formal education
(3) Non-White/Asiatic married man, 30 years old, with two years of formal education
(4) Non-White/Asiatic married man, 50 years old, with one year of formal education
(5) Non-White/Asiatic married man, 30 years old, with eight years of formal education
(6) White/Asiatic married man, 30 years old, with eight years of formal education

7. Conclusions

The main objective of this paper was to discuss the selectivity of migration and to make associations with rural poverty. In order to do so, mathematical simulations were done using the Roy and the human capital models as frameworks, which were composed of a long-term and a short-term equation. These simulations indicated that long distance flows, particularly the rural/urban ones might show a positive selection, mainly due to the short-term costs of migration.

The features discussed were the influence of social nets, age, regional low-skilled and high skilled workers wages, etc on the selectivity of migration.

Some aspects of these simulations were observed empirically in multinomial logistic models, such as the importance of age and schooling levels for the selectivity of migration. It was observed a general tendency of negative selection in rural/rural flows and a positive one in rural/urban and longer steps of migration. Two among the five Brazilian macroregions concentrate most of the very low-income migrant flows in Brazil, the North and Northeast ones, particularly the short distance migration with destiny in rural areas. These two facts indicated, as discussed in the mathematical simulations, that most low-income individuals are able to migrate, but only in short steps and low cost migration between rural areas.

In the North Region there are many rural areas with positive net migration (Golgher 2006b). As showed here, the short distance rural/rural migrants presents very low mean levels of schooling, they migrate from one locality to a similar close one, maybe with real chances of improving their economical situation, but with remarkable impact on deforestation of native vegetation (Laurence et al, 2001).

For the Northeast, at least till 2000, most areas had negative net migration (Golgher 2006b). As was observed in the mathematical and empirical models and simulations, low-income individuals show a lower propensity to migrate in long distance steps or rural/urban migration. Hence, the positive selection that may occur in areas with negative net migration might promote a vicious circle of negative feedback for economic and population regional aspects. Regional inequality may increase with this process, if positive aspects of emigration, such as remittances or knowledge transference are not significant (De Haan, 1999).

As proposed by De Haan (1999), most of the studies that analyzed rural and agricultural regional development did not give the appropriate importance to migration. Human mobility is much more common than normally assumed by the notion that population is essentially sedentary and would migrate only because of economical or environmental shocks. However, as was shown by Ghobadi et al (2005), migration is generally not an ex-post response to risks and shocks, but an ex-ante strategy of income and risks diversification, essential for the rural household. Therefore, given the importance of migration for the rural population, policies that promote mobility or, that increase the positive effects of migration, should be encouraged. Policies that diminish the costs of migration would have a positive impact on the range of possibilities for the low-income population strata. For instance, policies that: improve channels for information exchange; facilitate the absorption of the migrant in the destiny; minimize environmental damages; increase the effectiveness of the use of remittances for local development, etc (De Haan, 1999), are some of them.

REFERENCES

Barros, R., Henriques, R. and Mendonça, R. (2000). A estabilidade inaceitável: desigualdade e pobreza no Brasil. In: Henriques, R. (org.). Desigualdade e Pobreza no Brasil. Rio de Janeiro: IPEA.

Bell, P; Fisher, J; Baum, A.; Greene, T. Environmental psychology. Harcourt Brave Jovanovich College Publisher, 1990.

Borjas, G. (1987) Self-selection and the earnings of immigrants, America Economic Review 77 (4): 531-553.

——— (1996) Labor Economics. The McGraw-Hill Companies, Inc.

Cadwaller, M. Migration and residential mobility: macro and micro approaches. The University of Wisconsin Press, 1992.

Carey, F and Sundberg, R (1984) Advanced Organic Chemistry, part A, second edition, Plenum Press, New York and London.

Carvalho, J and Machado, C. (1992) Quesitos sobre migrações no Censo Demográfico de 1991. Rev. Bras. Estudos Pop., Campinas, v.9, n.1.

Castiglioni, A. (1989). Migration, urbanisation et dévelopement: le cas de l,Espírito Santo - Brésil. Ciaco Editeur.

Chiquiar, D and Hanson, G. (2002) International migration self-selection, and the distribution of wages: evidence from México and the United States, NBER working paper series, working paper n. 9242, http://www.nber.org/papers/ w9242.

De Haan, A. (1999). Livelihoods and poverty: the role of migration – a critical review of the migration literature. The journal of development studies, December, 36, 2, ABI/INFORM Global, pp 1.

Duarte, R. Migration and urban poverty in northeast Brazil (1979). Glasgow, University of Glasgow, (Ph.D. dissertation).

Ferreira, F, Lanjouw, P. and Neri, M. (2000). A new poverty profile for Brazil using PPV, PNAD and Census Data, Texto para discussão, No. 418, Departamento de Economia da PUC-RIO

FIBGE (2000) Censo Demográfico do Brasil de 2000.

Ghobadi, N., Koettl, J., Vakis, R. (2005) Moving out of poverty: migration insights from rural Afghanistan..www.mrrd.gov.af/vau/.

Golgher, A (2006a) Diagnóstico do processo migratório no Brasil 1: comparação entre não-migrantes e migrantes. Belo Horizonte, CEDEPLAR/FACE/UFMG, texto para discussão nº 282.

——— (2006b) Diagnóstico do processo migratório no Brasil 4: migração entre municípios. Belo Horizonte, CEDEPLAR/FACE/UFMG, texto para discussão nº 285.

Greenwood, M. Human Migration: Theory, Models, and Empirical Studies. Journal of Regional Science, v.25, p.521-544, 1985.

Gugler, J. Cities, poverty and development: urbanization in the third world. Oxford, Oxford University Press, 1992.

Haurin, D and Haurin, J. (1991) Youth migration in the United States: analysis of a deindustrializing region. In: Stillwell and Congdon: Migration models: macro and micro approaches. London and New York, Belnavn Press.

Hoffmann, R. (2000). Mensuração da desigualdade e da pobreza no Brasil. In: Henriques, R. (org.). Desigualdade e Pobreza no Brasil. Rio de Janeiro: IPEA.

Hollnsteiner-Racelis, M. (1998) Becoming an urbanite: the neighborhood as a learning environment. In: Gugler, J. The Urbanization of the Third World, Oxford, Oxford University Press.

IBRE/FGV (2005) Miséria em queda: mensuração, monitoramento e metas. Rio de Janeiro: Centro de Políticas Públicas do IBRE/FGV.

Kothari, U. (2002) Migration and Chronic Poverty, Chronic Poverty Research Centre, Working Paper No. 16, Institute for Development Policy and Management, University of Manchester.

Laurance, W., Cochrane, M., Bergen, S., Fearnside, P., Delamônica, P., Barber, C., D'Angelo, S. and Fernandes, T. (2001) The Future of the Brazilian Amazon**,** Science, Vol. 291, no. 5503.

Massey, D; Arango, J; Hugo, G; Koraoua, A; Pellegrino, A; Taylor, J. (1998) Worlds in motion: understanding international migration at the end of the millennium. Oxford, Clarendon Press.

Riani, J and Golgher, A. (2004) Indicadores educacionais confeccionados a partir de bases de dados do IBGE. In Introdução a Demografia da Educação, Rios-Neto and Ruas (org), Campinas, ABEP, pp 89-128.

Rigotti, I. and Carvalho, J.(1998) As migrações na grande região centro-leste. In: Encontro Nacional sobre migração, 1, Curitiba, 1997. Anais.., Curitiba: IPARDS/FNUAP, p.67-90.

Todaro, M. (1980) Internal migration in developing countries: a survey. In Easterlin, R. Population and economic change in developing countries. Chicago, University of Chicago Press for National Bureau of Economic Research.

Sandefur, G., Tuma, N. and Kephart, G (1991) Race, local labor markets, and migration, 1975-1983. In: Stillwell and Congdon. Migration models: macro and micro approaches. London and New York, Belnavn Press.

Stillwell, J. and Congdon, P. (1991). Migration modeling: concepts and contents. In: Stillwell and Congdon. Migration models: macro and micro approaches. London and New York, Belnavn Press.

Waddington, H and Sabates-Wheeler, R. (2003). How does poverty affect migration choice? A review of literature, Working Paper No T3, Development Research Centre on Migration, Globalization and Poverty, University of Sussex.