FACULTY WORKING PAPER 91-0115

Intermetropolitan Brazilian Migration: Estimates of a Multinomial Logistic Model

Ahmad Oran
Department of Economics

Bureau of Economic and Business Research
College of Commerce and Business Administration
University of Illinois at Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 91-0115

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

February 1991

Intermetropolitan Brazilian Migration: Estimates of a Multinomial Logistic Model

Ahmad Oran

Department of Economics

My thanks are due to the Brazilian Census Bureau (IBGE) for providing the data, and to the Center for Latin American and Caribbean Studies for awarding me Tinker Foundation research grant to undertake this study.

ABSTRACT

The one form of migration that has been extensively addressed is the rural-urban one while the other forms have not been sufficiently examined.

Data on the Brazilian metropolitan migration, based on the 1980 census, is employed to examine migrants' responses to the would-be socioeconomic migration determinants. Unlike other studies, this one considers the metropolitan area as the geographical unit over which migration is measured, and employs real values, rather than nominal ones, for the market variables as well. In addition, data are stratified by migrants educational attainments. Four groups are considered: no education, primary, secondary, and high. Finally, two structurally different migration variables are allowed for and studied by employing a multinomial logistic model.

INTERMETROPOLITAN BRAZILIAN MIGRATION: ESTIMATES OF A MULTINOMIAL LOGISTIC MODEL

I. Introduction

Over the last decades, there have been a number of studies on interregional migration. Several migration models were estimated on the basis of data from various Developed and Less Developed Countries (LDCs). Considerable evidence pertaining to this socioeconomic phenomenon in a large number of countries have been provided. Albeit we have greatly benefited from all works on internal migration, reservations have been kept concerning the way some migration issues, discussed below, were dealt with.

In keeping with the concern of researchers, this study intends to examine migration decision allowing for the debated issues. Several features add to the peculiarity of this study, among these: testing a unique set of data, considering other migration factors in addition to the traditional ones, testing a proposed migration variable that has never been tested before, stratifying data by migrants' educational attainment, and employing for the first time a multinomial logistic model to the Brazilian migration.

II. Debated Issues

It has been argued that explanatory variables are often poorly measured. One aspect that is noteworthy is the use of nominal values for the economic variables (market variables), while there is some evidence that real values (nominal ones deflated by a cost-of-living index) are more important for migration decision (Yap, 1977).

Another aspect that has been objected to is the level of aggregation, that is, the spatial unit over which migration is supposed to be measured. It has been argued that migration has not been appropriately measured due to the arbitrary choice of the level of aggregation[1] (Yap, 1977; Martine, 1987). State, for example, has been widely used as a spatial unit. The objection to this unit derives from the displayed degree of socioeconomic heterogeneity. In other words, within the same state one can identify various regions, thus various subgroups of population, showing different stages of socioeconomic conditions. In particular, state does not present a sufficient degree of economic integration to be employed as a proxy for classic labor market (Morgan, 1975/76). Consequently, differential migrants' responses, whether in terms of inmigration or outmigration, to migration factors can either be over- or underestimated.

Another debated issue is the specification of the migration variable. Pros and cons to the employed variables are found in the literature (Young, 1975; Vanderkam, 1976; Yap, 1977; Schultz, 1982; Shaw, 1985). One can easily verify that there is a general consensus on normalizing gross migration, i.e., transforming it into a migration rate; however, apparently, no consensus on how such normalization should take place has been reached yet. Various forms have been employed, yet the debate on the appropriate one seems to continue. Regrettably, there does not exist convincing statistical or economic reasons to support this or that normalization. It appears that the preferred procedure by most researchers is M_{ij}/P_i for it takes into consideration the population at the origin where the decision of moving was undertaken.

Young (1975) has suggested a different form of normalization, thus, a different migration variable. This proposed variable, which is an analogous one to the interregional trade parameter, takes on the following form: $(M_{ij}P)/(P_iP_j)$, where M_{ij} is the gross migration between regions i and j, P is the total population, and P_i and P_j are the population at origin and destination areas respectively. Several reasons were put forward by Young in support for this variable, these, however, are not discussed here. In spite of this, no one that I am aware of has tested this variable. Therefore, it is my intention to test Young's suggested variable against the most commonly used one, i.e., M_{ij}/P_i.

III. Type of Migration

Internal migration has largely been seen as synonymous with rural-urban migration for most studies concentrate on this form of migration. Although it has been recognized that a great deal of migration in LDCs occurs between rural and urban areas, urban-urban migration that is relatively great (becoming greater in some LDCs) in magnitude does not enjoy an equal recognition. It appears that it is not yet realized that in many LDCs rural-urban migration constitutes but a small part of all internal migration (Mazumdar, 1987).

Therefore, the type of migration under consideration in this study is the urban-urban form of it. This does not necessarily imply that other types of internal migration are not worth considering or that they need no more attention. What it means, however, is that urban-urban migration has not been sufficiently examined.

On the other hand, the demographic aggregation used in this study is the metropolitan area (MA). Until a reasonable spatial unit that would display, if not more homogeneous, less heterogenous socio-economic conditions is made available, this unit (MA) appears to be a better choice relative to state or any other arbitrary spatial unit.[2]

Internal migration is said to be a selective process where personal attributes play an important role in such selectivity. In order to shed some light on this issue as well as the whole migration process, migrants are studied allowing for one of their main attributes: educational attainment. Hence, migrants are categorized into four educational groups: none, primary, secondary, and high.

IV. Explanatory Variables

It is quite safe to say that there is no unique set of variables to be applied to all and every migration study. Various different sets of variables have been used in different works. In fact, it cannot be otherwise for researchers may differ on the relevance of some factors to migration, be constrained by the available data, and/or attempt to emphasize different aspects of migration. Moreover, the choice of explanatory variables is highly affected by the level of development of the area or country under consideration. While a migrant of a poor country, for example, struggles for a job, a migrant of a rich country is most likely to be more concerned with particular aspects of that job.[3]

In any case, it is agreed upon that migratory selective nature takes into account that migration decisions are not only influenced by regional and personal characteristics, but by the difficulty of the process of moving itself as well. Although origin and destination attributes that are commonly considered to be the most affecting the attractiveness of migrant have largely been suggested in the literature, the way by which these attributes are measured varies considerably.

Quite surprisingly, it seems that the push factors that explain outmigration are not primarily the economic conditions of the considered area but rather the demographic attributes of the population of that area. The pull factors, the ones that demarcate the migrant's choice of where to move, appear to be primarily the economic conditions of the other areas (Hoover et al., 1984).

V. Data and Variables

Most migration studies, mainly in LDCs, have been limited by the available data in one way or another. Among other reasons for choosing Brazil as a case of study is the quality of the available data. Indeed, the Brazilian census shows continuous improvement since 1940 (DaCunha, 1986; Martine, 1987). As an example of this, the 1980 census, on which this study is based, presents further improvements.

The explanatory variables considered in this work, the reasons behind their choice, and their expected effect on migration process are presented in turn.

5.1 Earnings--RER

Since not all migrants are necessarily monthly-paid workers, the average monthly wage is considered for wage-earners and the average monthly income is considered for the non-wage-earners, self-employed. On the other hand, it seems that employing real values, as it is the case of this work, rather than nominal ones is a much better way for measuring the effect of this variable on the process of migrating (Yap, 1977; Hoover et al., 1984). Typically, migration is expected to flow from a low real earning metropolitan area to a high real earning one, *ceteris paribus*.

5.2 Employment--REM

In addition to introducing into the model the labor market conditions at origin and destination areas, the employment variable represents for a potential migrant the probability of finding a job given migrant's personal attributes. It is, *ceteris paribus*, expected that the flow of migration be directed toward metropolitan areas that display high rate of employment.

5.3 Per Capita Income--RPC

This variable may be used as a substitute for the earning one since per capita income may better display interregional income differentials than income or wage rate does. However, it is not easy beforehand to predict whether this variable or the earning one fits best the model. Irregardless of this, migrant's behavior toward this variable is expected to be the same as the one toward the earning variable.

5.4 Density--RDN

Density is said to be partially affected by migrants. If potential migrants are able to obtain information about the destination areas (mainly about the labor market conditions) and some other kind of aid upon arrival, from relatives and friends, density can then be seen as a proxy of a "snowball" effect of the previous migration (migration stock) on the studied migration variable (Sahota, 1968). Nonetheless, it may not become a significant proxy variable for the previous migration.

Moreover, Mazumdar (1987) argues that population density, in the origin area, may have an indirect effect upon migration through its influence on transportation cost, marginal product of labor, etc. One may add that population density could also exert such effect in the destination areas constituting a deterrent migration factor. Consequently, depending on how density is interpreted by migrant, it may have either push or pull effect on migration. Therefore, it is also one of these variables that cannot be beforehand predicted.

5.5 Urbanization--RUR

Although this work is concerned with already urbanized areas (both at origins and destinations), different levels of urbanization are likely to exist among urban areas themselves. Hence, it is reasonable to expect the existence of substantial differences among metropolitan areas not only in terms of the labor market conditions, but also in terms of the educational, intellectual, and other amenities, including the so-called "bright-lights." It is, therefore, expected that migrants would be more attracted to a high degree urbanized area, *ceteris paribus*.

5.6 Government Spending--RGS

There is no doubt that public utilities, such as health service, power, garbage collection, public transportation, water supply, etc. are largely regarded important regional attributes by the general public, including migrants. Consequently, it seems very appealing to include a variable that accounts for the effects of the types and levels of public services on the migration decision of where to locate. Therefore, migrants are expected to be attracted to areas with high government spending, *ceteris paribus*.

5.7 Schooling--RSC

The same stratification applied to migrants' educational attainment (none, primary, secondary, and high) is also used to the population at large. Hence, the percentage of the population that pertains to each educational group at both origin and destination areas determines the content of this variable.

Different, and possibly conflicting, hypotheses can be constructed about the effect that this variable might exert on migration decision. Schooling, as defined here, may turn out to be either pull or push factor depending on whether migrant is risk loving or risk averse. Although it is believed that schooling influences migration decision, a clear pattern in terms of migrants' response to it cannot easily be beforehand predicted.

5.8 Housing Rent--RHR

Housing rent constitutes, if not the largest, one of the most significant expenditure item on any consumer's list, including migrant.[4] From the point of view of urban theory, this variable appears to be the candidate proxy variable for the host of amenities. As such, there are two conflicting sign expectation of housing rent on migration decision (Graves, 1983).

Given consumers' large expenditure on housing rent, I find it very appealing to use this variable as an explanatory variable, for it presents itself as a potential candidate for a major migration deterrent factor. Hence, migration flows are expected to be directed towards metropolitan areas that display low housing rent, *ceteris paribus*.

5.9 Geographic Distance--Dist

Distance has traditionally been the only variable that is expected to account for the entire migration cost (moving cost, opportunity cost, and psychic cost). Truly, distance per se seems to be loosely related to migration for costs involved in migration, as noted, are more than moving cost. Uncertainty, risk and investment of time, restraints on migration per se and what is sometimes referred to as social distance (difficulty with social adjustment after arrival) are significant parts of migration cost (Schwartz, 1973; Hoover et al., 1984).

Given the subjective character of the large part of migration costs and the consequent difficulty of quantifying them, distance remains the major proxy variable for migration costs.

In addition, it has been argued that current migrants are affected by previous ones, stock of migration, because they comprise an important source of information. Thus, migration studies should pay attention to previous flows of migration and try to account for their effect on the migration decision. However, stock of migration, likewise distance, is a link between places. Hence, the omission of one would pragmatically increase the importance of the other. If previous flows of migration are omitted, the apparent importance of distance would increase. For if previous migration were dominantly inversely related to distance, high stock of migrants would be found in destination close to an origin. Therefore, the information passed along by previous migrants would counteract the effect exerted by distance. By the same token, distant destination would have a smaller number of migrants because of distance and the small stock of relatives and/or friends.

The apparent effect of distance on migration would be upwardly biased when the effect of migration stock is not allowed for (Mueller, 1982). All in all, the effect of previous migration is believed to be accounted for, though implicitly, via the distance variable.

VI. Economic and Statistical Framework

Interregional migration has been approached from general and partial equilibrium framework as well. Within the latter approach, internal migration has been considered along two, not mutually exclusive, lines: investment and consumption.[5] As an investment decision migrants are viewed as profit maximizing units. Migrants, as labor force suppliers, are willing to supply their labor as long as the outcome of such transaction is maximized. In other words, a migration decision is undertaken if the present value of all streams of benefits net of associated costs is maximized (Hicks, 1932; Sjaastad, 1962).

On the other hand, since a migration decision implies a move from one place, called origin, to another place, called destination, the formulation of the migration function should specify whether or not conditions at both areas (origin and destination) are expected to equally affect the migration variable. If an equal (symmetric) effect is hypothesized, as is the case in this study, it will be assumed that origin and destination characteristics exert the same effect on the migration variable but in opposite directions. This amounts to assuming perfect information in the labor market. Accordingly, the migration function can, in its general form, be formulated as follows

$$M_{ij} = f(X_j X_i^{-1}; D_{ij}) \tag{1}$$

where M_{ij} is the gross migration from i to j, X_i and X_j are the regional attributes at origin and destination areas respectively, and D_{ij} is the distance between i and j.

To incorporate the economic theoretical understanding of migration, researchers have been employing the multinomial (polytomous) logistic (logit) model. Since its introduction into economics by McFadden (1974) and then into migration studies, this model has become the main statistical tool for migration analysis substituting the long time undisputed gravity model.

This specification, multinomial logistic model, centers around the assumption that migration decision is a choice to be made over a finite set of different and mutually exclusive discrete alternatives. Accordingly, a migrant is assumed to face a limited number of alternative locations for his choice including his place of origin. The basic formulation of the model as found in the literature (Schultz, 1982; Fields, 1982) is

$$P_{ij} = \exp Z_{ij} / \Sigma_{j=1}^{n} \exp Z_{ij} \qquad i,j = 1,\ldots,n \tag{2}$$

and for every region of origin

$$\Sigma_{j=1}^{n} P_{ij} = 1 \tag{3}$$

The selection probabilities (P_{ij}) that an individual migrates to a place j being in i, is assumed to depend on a vector of weighted personal and regional attributes Z_{ij}. The functional specification of this vector has been suggested by Schultz (1982) to be a linear function in natural logarithms of the characteristics of the origin (X_i) and destination (X_j) regions and distance (D_{ij}).

Therefore, the Z_{ij} function can be formulated as

$$Z_{ij} = \alpha + \Sigma_{k=1}^{K} \beta_k \ln X_{ki} + \Sigma_{k=1}^{K} \delta_k \ln X_{kj} + \gamma \ln D_{ij}$$

$$i,j = 1,\ldots,n; \; k = 1,\ldots,K \tag{4}$$

Moreover, the constraint, equation (3), imposed on the odds ratio, equation (2), can be conveniently accounted for by combining these two equations. That is, by expressing the migration probabilities as ratios and taking the log of these ratios, i.e., $\ln(P_{ij}/P_{ii})$.

Allowing for the assumed symmetrical effect and following the assumption made by Fields (1982) that P_{ii} can be considered as a constant across labor markets, then the formulation of the logit model can be approximated by

$$\ell n\ P_{ij} = \alpha + \Sigma_{k=1}^{K} \beta_k\ \ell n\ X_{ki} + \sum_{k=1}^{K} \delta_k\ \ell n\ X_{kj} + \gamma\ \ell n\ D_{ij} \qquad (5)$$

Since the full application of the multinomial logistic model, $\ell n(P_{ij}/P_{ii})$, is ruled out given the demanding nature in terms of data and computer work, the left-hand side of equation (5) is substituted by a proxy variable that allows us to introduce gross migration to be empirically studied. The specification that has been widely used is the one shown in the following general formulation of the logistic model

$$\ell n(\frac{M_{ij}}{P_i}) = f[\ell n(X_j X_i^{-1}),\ \ell n\ D_{ij}] + V_{ij} \qquad (6)$$

VII. Application and Empirical Results

Empirical studies have always been bounded by the availability of appropriate data. Although the data at hand concerning the Brazilian intermetropolitan migration--specially tabulated and provided by the Brazilian Census Bureau (IBGE)--are of a high quality, they fell short of providing the needed information for the following two variables.

Information on the explanatory variable "Government Spending" by metropolitan areas is not available at this level of aggregation. The value employed for this variable is the one calculated for the capitals of the correspondent metro areas. This information was obtained from a survey conducted by IBGE called Dados basicos.

Cost-of-living index is also not computed per metro area. Therefore, the value employed to transform the nominal values of the market variables into real ones is the value of this index calculated for the capitals of the metropolitan areas.

Turning to the application itself, the general multinomial logistic model estimated in this study is

$$\mathrm{Ln}\left(\frac{M_{ij}}{P_i}\right) = f(\mathrm{RER},\ \mathrm{RSC},\ \mathrm{REM},\ \mathrm{RDN},\ \mathrm{RUR},\ \mathrm{POP}_j,\ \mathrm{Dist}) + V_{ij} \qquad (7)$$

where the left-hand side of (7) is the dependent variable or the rate of migration. The right-hand side of (7) shows the employed explanatory variables namely: real earning (RER), schooling (RSC), employment rate (REM), density (RDN), urbanization (RUR), population at destination (POP_j), and geographical distance (Dist), in addition to the disturbance terms (V_{ij}).[6]

It should be noted that the model shown in (7) is the reduced form of the originally intended one. Some variables, housing rent (RHR) and government spending (RGS), had to be excluded from the model due to the existence of multicollinearity between both variables, on one hand, and earning and density, on the other, as inferred from the strong simple correlation coefficient among them.[7] Multicollinearity may result in unstable estimates from sample to sample, therefore, unreliable to be useful for the analyses. Since good information on a limited set of relationships is far better than poor information on a large one, these variables were eliminated from the model.

The maximum number of cross-migration flows among the Brazilian nine metropolitan areas per type, sex, and educational attainment is

81 (9x9) pairs of migration flows. However, 72 observations only are considered in the empirical work, i.e., the main diagonal is excluded.

7.1 Regression Analyses

This section presents the results of fitting the place-to-place migration model to four subsamples. That is, the full sample is stratified by migrants' educational attainments (none, primary, secondary, and high). Each subsample consists of 288 (72x2x2) observations for the nine metropolitan areas.

Four different functional forms (models) are used to test migrants' behavior. The result of these tests are presented in Tables 3-6. In addition to substituting per capita income (RPC) for real earning (RER), two dependent variables are considered. Models 1 and 2 employ the migration rate specified by (M_{ij}/P_i) while Models 3 and 4 regress the migration rate specified by $(M_{ij}P/P_iP_j)$ on the explanatory variables. Before turning to the empirical results, the reader is reminded that estimates are in the log of the ratio of regional attributes in origin and destination. Therefore, elasticities close to zero suggest the existence of a symmetrical relationship between variables in question and vice versa.

Real earning (RER) is statistically significant for all migrants, except in the case of migrants of primary education, and displays the expected, positive, sign in all models. The size of the estimates suggest an asymmetrical relationship between the values of real earning at origin and destination.

When per capita income (RPC) is substituted for real earning, one verifies the same behavior as in the case of real earning, i.e., statistically significant and displaying the expected sign in all cases. In general, the elasticities in this case show larger size and t-ratios than in the real earning case. This behavior appears to suggest that per capita income is a better explanatory variable. Interestingly enough, the size of elasticities is inversely related to migrants' educational attainments. Larger estimates are verified for migrants of no formal education, and smaller ones for migrants of high education. This strongly suggests that income is seen by none educated migrants as a crucial determinant in their decision to migrate. Although high educated migrants recognize the importance of income, they do not consider it to be a decisive one.

Schooling is statistically significant, in all models, in the case of migrants of secondary and high education while significant in only four models out of eight in the case of migrants of none and primary education. The sign displayed by schooling across migrants educational groups confirms what I have said earlier, i.e., schooling can be differently seen by migrants. This variable does not appear to be a deterrent to none and high educated groups contrary to primary and secondary groups. It is quite difficult to provide a reasonable interpretation for this behavior. For any possible interpretation depends on how each educational group evaluates the role of schooling. That is, if it means higher future gains and better educational opportunities, then the response of none and high educated groups seems to be reasonable. On the other hand, if primary and secondary groups view schooling as a sign of a potentially very competitive market, their responses seem also to be reasonable.

Likewise, employment rate presents very interesting responses. This variable displays statistically significant estimates for all groups, except for migrants of primary education, although with the wrong (negative) sign. Employment rate is considered, according to these empirical results, as a deterrent to migration to three groups of migrants while the other group does not respond to it at all!

Studying the interstate Venezuelan migration, Schultz (1982) found employment rate at origin never statistically significant related to migration at the 0.05 level, and destination condition are positive and statistically "convincing" only for higher educated migrants. Moreover, he indicated that low educated migrants present no response to employment in destination. Studying recent and lifetime Brazilian urban migration, Oran's (1990) findings show employment variable displaying statistically significant with the wrong sign for highly educated recent and lifetime migrants. Moreover, no statistically significant estimates were verified for lowly educated migrant in that work. Field's study (1982) of the Colombian migration indicates that this variable presented wrong signs more often than not, which led him to eliminate employment from his model. As can be noticed, all these works have considered Latin American migrations. Interestingly enough, all of them have some difficulties with employment rate one way or another.

Urbanization presents statistically significant positive (expected) sign in one or more models for each educational group. The size of estimates, across migrants' groups, and consequently the importance of this variable to migrants, is inversely related to educational level. That is, results strongly suggest a decrease in the importance of this variable to migrants as their educational attainments increase. In fact, the elasticity of urbanization in the case of highly educated migrants is statistically insignificant but in one model only. Even in this case, the null hypothesis is almost accepted at the 0.05 level. If migrants see in urbanization an indication of the availability of recreational, educational, entertainment facilities including what is termed "bright lights," then the presented responses, seem to be consistent.

Density displays statistically significant elasticities for all migrants in all models with relatively high t-ratios indicating its strong relationship to migration. This variable seems to be a push factor due to population pressure, or a pull factor if it is taken to represent a "snowball" effect of migration stock. The size of the estimates for all educational groups in all models are very close from each other showing that density is seen equally by all migrants.

Distance has the same deterrent effect on migration as found in many other studies of this genre. It displays negative and statistically significant estimates for all migrants in all models. Moreover, the size of the estimates is inversely related to migrants' educational attainments. That is, the highest elasticity is verified for migrants of no educational attainment, then it decreases as the level of educational attainment increases to reach the lowest estimate for the highly educated migrants. This implies that distance although considered deterrent by all migrants, it becomes less deterrent as migrant's educational attainments increase. The performance of this variable corresponds exactly to what was expected.

On the other hand, the explained proportions of variance are all above 50 percent. Interestingly enough, R^2 in all models across educational groups increases as migrants' educational attainment increases, establishing a direct relationship between them.

The explanatory power of the employed models, as measured by the F-values, displays not only significant but also interesting results. Across migrants, the explanatory power shows a direct relationship between the size of the F-value and migrants' educational attainment. That is, in all models, the explanatory power increases as migrants' educational attainment increases. This suggests that the higher migrant's education is, the better one can explain his behavior. In short, the goodness-of-fit of models varies directly with education.

In comparing the employed migration variables, (M_{ij}/P_i) versus $(M_{ij}P/P_iP_j)$, across models and migrants, similar migrants' responses were obtained. However, the interesting results are verified when the explanatory power of models are compared. In no model, across migrants, M_{ij}/P_i presents larger F-value than of the other variable. Indeed, Young's suggested variable presents better goodness-of-fit in all models. Therefore, all in all, I would say, based on the results of this study, that either variable can be employed although I am slightly inclined to accept Young's variable provided that similar results are obtained in other studies.

7.2 Tabular Evidence

The basic data for the Brazilian metropolitan areas by migrants' educational attainment in Table 1 allow us to make the following observations:

(a) When a migrant is defined to be someone who lives in a metropolitan area different from the one in which the migrant was born or lived in before, the rates of inmigration across the metro area range from 2.33 to 6.29; 3.54 to 7.14; 3.04 to 5.62; and 2.31 to 5.00 for migrants of none, primary, secondary, and high educational attainment, respectively. The lowest inmigration rate, in all cases, is associated with Belem metro area; except in the case of highly educated migrant, Fortaleza metro area displays the lowest rate. The highest inmigration rates are associated with Sao Paulo (none and primary educated migrants) while in the other two cases, the highest inmigration rates are associated with Rio de Janeiro metro area.

(b) Average real earning, across metro areas, shows significantly large difference between the lowest and highest earning for each group of migration. These differences are 100%, 87%, 83%, and 44% starting with educational attainment none. It is to be observed that differences between real earning, while they show regional disequilibria, decrease as educational attainment increases. However, even in the case of highly educated migrants, the gap remains very wide (44%).

(c) Average inmigration rates across metro areas for all educational groups, except high educational one, show direct relationship with the average employment rate. Contrary to this is the behavior of migrants of high educational attainment. In this case, the highest employment rate and the lowest inmigration rate are associated with each other and vice versa. In fact, in the case of highly educated migrants, results strongly suggest an inverse relationship across areas between these two variables. This behavior becomes more difficult to interpret if one considers at the same time the real earning rates associated with these areas. While inmigration rate for Rio de Janeiro metro area, for instance, is 116% greater than that of Fortaleza metro area, real earning differential is 18% favorable to the former area. Therefore, any plausible explanation cannot be based only on earning factor. Although one can advance several reasons to justify this behavior, such as lack of information on other areas, the importance of other regional attributes, the risk loving nature of migrants, etc., a full interpretation requires a thorough investigation.

Turning to the basic data for the regional attributes of the Brazilian metro areas (Table 2), the following observations can be made:

(a) Average inmigration rate by metro area ranges from 2.83 to 6.00 for Belem and Sao Paulo metro areas, respectively.

(b) Average geographical distance ranges from 1391 to 2422 km.

(c) Average government spending ranges from 7,074 million cruzeiros to 298,230 million cruzeiros, where the latter is larger than the former by a factor of 42 to one!

(d) Average real housing rent ranges from 35.51 to 39.90 cruzeiros.

(e) Low inmigration rates are associated with high distance, low government spending areas, and low housing rent area, except in the case of Belem metro area. Moreover, high inmigration rates are associated with low distance, high government spending areas, and high housing rent area. All in all, migrants' behavior represented by their inmigration rates is consistent vis-a-vis distance and government spending. However, there is a strong evidence that housing rent is a consequence of migration rather than a cause of it, not confirming, therfore, the expected behavior. Nonetheless, the performance of this variable may have been affected by the use of cost-of-living index since housing rent constitutes part of this index. Consequently, by using real values for the market variables, it seems that one is implicitly allowing for housing rent.

VIII. Conclusions

The main intent of this paper is to study intermetropolitan Brazilian migration, by employing a multinomial logistic model, so that some light on the decision for migrating can be shed allowing for migrant educational attainment.

To do so, several explanatory variables are considered where the market ones are measured in real terms. Estimates are obtained by examining four different models that test two different migration variables. The final considerations based on the statistical results in this work and some other basic data are summarized in turn.

Not surprisingly, real earning and/or per capita income constitute a major migration pull factor. Across educational groups and metro areas, migrants are attracted to areas that display high real earning. It is worth observing the existence of a significant earning differential not only across metro areas but also across educational groups that strongly suggests regional disequilibria. Employment rate turned out to be a migration deterrent according to the regression analysis, yet its importance to migration decision is detected when considering the basic data. Considering real earning and employment variables, real earning appears to have more weight in migrants' eyes. Such observation is clearly verified in the case of migrants of high education where these are attracted to area displaying not only high earning but low employment rate as well. Across educational groups and metro areas, two regions have been distinguished in their extremely high inmigration rates, Sao Paulo and Rio de Janeiro, which reinforces the observed regional disequilibria.

Turning to the regional attributes, distance confirms its strong deterrent nature to migration. Although density displays signs of pull factor, urbanization shows higher degree of attractiveness. On the other hand, the none traditionally tested regional factors, housing rent and government spending, while the latter confirms its importance to migration given the verified direct relationship of this with inmigration rate, the same cannot be said about the former, where this appears to be an effect of migration rather than a cause of it.

Finally, there is a strong statistical evidence to support the use of the alternative migration variable tested in this study.

NOTES

[1]An appropriate spatial unit is the one that is taken to represent a single labor market. In other words, it would be an area that is well defined occupationally, industrially, and geographically such as workers are willing to change jobs in a relatively free manner (see Kerr, 1950).

[2]In 1988, two Brazilian researchers of the Centro de Desenvolvimento e Planejamento Regional--CEDEPLAR were studying a new spatial division. According to this study, Brazil may become divided into approximately 55 "homogeneous" areas.

[3]Shaw (1985) outlines various sets of migration factors that vary relative to the level of development of a country.

[4]According to the Brazilian Estudo Nacional da Despesa Familiar (National study of family expenditure) of 1974/75 in the state of Sao Paulo, housing rent (including related urban taxes and service fees) was found to be the most significant individual component of a household expenditure. In the city of Sao Paulo, housing rent is found to be almost one-quarter of households' expenditure, or 23.7% (Vieira, 1984).

[5]In fact, the general framework is utility maximization where investment approach is a special case. For this is also a utility maximization that equates marginal utility of leisure to zero. See Mueller (1982) for utility maximization approach.

[6]The disturbance terms (V_{ij}) are assumed to be independently and identically distributed random variables. Each term of these is of a zero mean and constant variance across observations.

[7]Simultaneity bias arises if the relationship between two variables cannot be assumed to be in the same direction. In fact, most studies of place-to-place migration suffer to some extent from this problem. One way of minimizing simultaneous bias is to employ a system of simultaneous equations rather than employing a single equation model.

REFERENCES

DaCunha, Jose M. P. (1986). Aspectos quantitativos do processo de redistribuicao espacial da populacas no Estado de Sao Paulo--1960/80. Anais do V Encontro Nacional de Estudos Populacionais, V2, ABEP.

Fields, Gary S. (1982). Place-to-place migration in Colombia. Econ. Development and Cultural Change 30(3), 539-558.

Graves, P. E. (1983). Migration with a composite amenity: The role of rents. Journal of Regional Science 23, 541-45.

Hicks, J. R. (1932). The Theory of Wages. London: Macmillan Press.

Hoover, Edgar and F. Giarratani (1984). An Introduction to Regional Economics. New York, NY, Alfred A. Knopf, Inc.

Kerr, C. (1950). Labor markets: Their character and consequences. American Economic Review 40, 278-95.

Martine, George (1987). Internal migration in Brazil. Preliminary draft.

Mazumdar, Dipak (1987). Rural-Urban Migration in Developing Countries. World Bank Reprint Series: number 422.

McFadden, Daniel (1974). Analysis of qualitative choice behavior. In P. Zarembka, ed., Frontier in Economics, New York: Academic Press.

Morgan, C. (1975/76). Is out-migration affected by economic conditions? Comment. Southern Economic Journal 42, 752-58.

Mueller, Charles F. (1982). The Economics of Labor Migration: A Behavioral Analysis. New York, Academic Press.

Oran, Ahmad (1990). Recent versus Lifetime Brazilian Intermetropolitan Migration: Estimates of a Multinomial Logistic Model. Doctorate Dissertation, University of Illinois at Urbana-Champaign, 1990.

Sahota, Gian S. (1968). An economic analysis of internal migration in Brazil. Journal of Political Economy 76 (March/April), 218-245.

Schultz, T. P. (1982). Lifetime migration within educational strata in Venezuela: Estimates of a logistic model. Econ. Development and Cultural Change 30(3), 559-593.

Schwartz, Aba (1973). Interpreting the effect of distance on migration. Journal of Political Economy 81(5), 1153-1169.

Shaw, R. Paul (1985). Intermetropolitan Migration in Canada. Toronto: NC Press Ltd.

Sjaastad, Larry (1962). The cost and returns of human migration. Journal of Political Economy, Supplement 70, 80-93.

Vanderkamp, J. (1976). The role of population size in migration studies. Canadian Journal of Economics 9, 508-17.

Vieira, Claudio A. (1984). Urbanizacao e Custo de Reproducas da Forca de Trabalho. Series Ensios Economicos, V, 38, Sao Paulo: IBE/USP.

Yap, Loren (1977). Attraction of cities--A review of the migration literature. Journal of Development Economics 4(3), 239-264.

Young, Geoffrey (1975). The choice of dependent variable for cross-section studies of migration. Canadian Journal of Economics 8(1), 93-100.

Table 1

Basic Data for the Brazilian Metropolitan
Areas by Educational Attainments (N=36)

	Migrants' Educational Attainments											
	None			Primary			Secondary			High		
Destination Areas	Inmig. Rate**	ER*	EM %	Inmig. Rate**	ER*	EM %	Inmig. Rate**	ER*	EM %	Inmig. Rate**	ER*	EM %
Belem	2.33	16.74	36.52	3.54	27.33	45.89	3.04	63.63	65.18	2.41	182.05	89.93
Fortaleza	2.81	16.25	44.67	3.90	28.30	47.24	3.29	71.24	69.16	2.31	199.36	90.55
Receife	3.09	13.78	38.33	4.26	25.37	45.45	3.83	66.54	68.04	3.00	193.56	89.18
Salvador	3.44	19.61	42.81	4.29	39.73	47.02	3.81	92.29	71.00	3.37	241.64	88.75
Belo Horizonte	3.31	18.08	41.47	4.53	34.73	53.27	3.64	87.06	71.60	2.87	229.40	88.64
Rio de Janeiro	5.70	21.18	39.60	6.61	37.03	49.39	5.62	94.20	68.15	5.00	235.65	84.81
Sao Paulo	6.29	27.76	45.19	7.14	47.37	57.44	5.60	116.25	71.64	4.94	262.82	85.92
Curitiba	2.74	18.60	43.13	4.16	33.90	52.70	3.42	77.38	68.22	2.71	189.21	85.68
Porto Alegre	2.44	23.00	44.88	3.89	37.84	56.39	3.10	84.42	72.93	2.58	210.99	88.50

*Values in cruzeiros, 1980.
**Inmigration rate is in the log form.

Note: Variables are real earning (ER), and employment rate (EM), all in terms of destination areas.

Source: Brazilian Census, 1980.

Table 2

Basic Data for the Brazilian Metropolitan Areas
by Regional Attributes (N=144)

Destination Areas	Inmig. Rate**	Distance (KM)	Govt.* Spending	Housing* Rent
Belem	2.83	2422	7,074	37.15
Fortaleza	3.08	2222	14,548	36.99
Receife	3.55	1958	20,659	35.51
Salvador	3.73	1643	23,288	37.89
Belo Horizonte	3.59	1391	28,731	35.62
Rio de Janeiro	5.70	1479	137,702	38.40
Sao Paulo	6.00	1468	298,230	39.90
Curitiba	3.26	1686	17,980	36.19
Porto Alegre	3.00	2228	26,019	35.51

*Values in cruzeino (1980).
**Inmigration rate is in the log form.

Source: Oran (1990).

Table 3

Estimates of the Multinomial Logistic Model for Intermetropolitan Brazilian Migration
Educational Group: None
(N=288)

Variables	Model 1 Estimates	Model 2 Estimates	Model 3 Estimates	Model 4 Estimates
Constant	2.3760	0.8139	5.8659	5.8646
Dist	-1.0292*	-0.9487*	-0.8482*	-0.8481*
	(-7.407)	(-7.083)	(-6.306)	(-6.633)
POP_i	0.3907*	0.3886*		
	(2.818)	(3.048)		
POP_j	0.9296*	1.0780*		
	(8.054)	(9.530)		
RSC	0.1384	2.6059*	0.5316	3.4763*
	(0.216)	(3.400)	(0.794)	(4.520)
REM	-2.5922	-2.6470*	-2.9028*	-3.6436*
	(-1.945)	(-2.563)	(-2.101)	(-3.467)
RER	1.5436*		1.0348	
	(2.485)		(1.698)	
RPC		3.4831*		3.4241*
		(5.676)		(5.760)
RDN	0.5822*	0.6327*	0.8476*	0.7735*
	(4.763)	(6.046)	(7.236)	(7.762)
RUR	8.8920*	3.7686	13.092*	7.3513*
	(4.490)	(1.820)	(6.860)	(3.616)
R^2	0.53	0.57	0.50	0.54
F-value	39.15	45.91	47.05	57.07

*Estimates significant at the 0.05 level.

Note: Values in parentheses are the t-ratio values.

Table 4

Estimates of the Multinomial Logistic Model for Intermetropolitan Brazilian Migration Educational Group: Primary (N=288)

Variables	Model 1 Estimates	Model 2 Estimates	Model 3 Estimates	Model 4 Estimates
Constant	-1.3827	-1.5861	5.3268	5.4200
Dist	-0.9262*	-0.8798*	-0.7651*	-0.7767*
	(-7.989)	(-7.770)	(-7.103)	(-7.486)
POP_i	1.0089*	0.8319*		
	(6.029)	(5.070)		
POP_j	0.6715*	0.8406*		
	(6.212)	(7.542)		
RSC	1.1577*	-1.0066	1.4033*	-1.1258
	(2.381)	(-1.407)	(2.903)	(-1.585)
REM	0.1514	0.0980	0.0663	0.0462
	(0.724)	(0.493)	(0.314)	(0.233)
RER	0.4286		0.6891	
	(1.047)		(1.781)	
RPC		2.0988*		2.2778*
		(4.143)		(5.044)
RDN	0.6289*	0.5995*	0.7404*	0.7099*
	(7.117)	(7.000)	(10.123)	(10.051)
RUR	6.3354*	1.8184	7.3791*	2.5743
	(4.083)	(1.033)	(5.031)	(1.550)
R^2	0.56	0.58	0.54	0.57
F-value	43.68	48.19	54.63	62.58

*Estimates significant at the 0.05 level.

Note: Values in parentheses are the t-ratio values.

Table 5

Estimates of the Multinomial Logistic Model for
Intermetropolitan Brazilian Migration
Educational Group: Secondary
(N=288)

Variables	Model 1 Estimates	Model 2 Estimates	Model 3 Estimates	Model 4 Estimates
Constant	-4.4865	-4.9928	5.7012	5.7012
Dist	-0.8382*	-0.8390*	-0.8046*	-0.8046*
	(-9.603)	(-9.618)	(-9.919)	(-9.801)
POP_i	1.5247*	1.6678*		
	(9.255)	(10.304)		
POP_j	0.6607*	0.5916*		
	(5.149)	(4.773)		
RSC	-1.3119*	-1.8004*	-1.2341*	-1.6788*
	(-2.411)	(-3.191)	(-2.253)	(-2.944)
REM	-6.9665*	-6.4831*	-7.6330*	-7.0252*
	(-5.294)	(-5.025)	(-6.107)	(-5.616)
RER	2.2987*		2.6276*	
	(6.695)		(8.493)	
RPC		1.4425*		1.5940*
		(6.710)		(7.983)
RDN	0.4276*	0.5974*	0.3396*	0.5119*
	(6.114)	(9.237)	(5.395)	(8.941)
RUR	4.5767*	6.4648*	4.0819*	6.3248*
	(3.484)	(5.350)	(3.160)	(5.364)
R^2	0.67	0.67	0.62	0.61
F-value	72.37	72.44	76.26	73.34

*Estimates significant at the 0.05 level.

Note: Values in parentheses are the t-ratio values.

Table 6

Estimates of the Multinomial Logistic Model for Intermetropolitan Brazilian Migration Educational Group: High (N=288)

Variables	Model 1 Estimates	Model 2 Estimates	Model 3 Estimates	Model 4 Estimates
Constant	-4.2093	-4.1995	4.1166	4.1220
Dist	-0.6433* (-7.758)	-0.6491* (-7.800)	-0.5860* (-7.827)	-0.5867* (-7.749)
POP_i	1.4478* (9.267)	1.4917* (9.615)		
POP_j	0.5796* (5.365)	0.5399* (5.082)		
RSC	1.2283* (5.015)	1.0256* (3.180)	1.1806* (5.157)	0.9770* (3.100)
REM	-8.3111* (-4.623)	-5.9338* (-3.895)	-10.7817* (-6.578)	-7.7975* (-5.436)
RER	1.4273* (2.297)		1.9484* (3.295)	
RPC		0.5192 (1.720)		0.6175* (2.090)
RDN	0.5058* (5.888)	0.6406* (10.602)	0.4487* (5.833)	0.6327* (12.444)
RUR	0.4448 (0.354)	1.7833 (1.837)	-0.2482 (-0.207)	1.7215* (1.834)
R^2	0.70	0.70	0.69	0.68
F-value	81.96	81.00	102.35	99.03

*Estimates significant at the 0.05 level.

Note: Values in parentheses are the t-ratio values.