RURAL-URBAN MIGRATION AND SELF-SELECTION IN TUNISIA[1]

Ben Jelili Riadh*
Université de Bretagne Sud, C.E.I.E.
Riad.ben-jelili@univ-ubs.fr.

Mzali Hassen
Faculté de Sciences Economiques et de Gestion de Tunis.

Abstract : With its theoretic robustness and intuitive appeal, the human capital model of labor migration has been successful in explaining several empirical regularities of the migration process. While adhering to a similar approach, the purpose of this paper is to estimate, using survey data from Tunisia, a model of returns to rural-urban migration which accounts for self-selection of migrants.

Of particular interest in this paper is the sign and significance of the selectivity terms in the movers and the stayers group.

The results lead to the conclusion that by purely statistical assessment the expected monetary gains effect is significantly different from zero, but that by economic considerations it is small. This low effect can be explained by the omission of other relevant variables from the analysis of rural-urban migration in Tunisia.

In the migrant-earnings equation the selectivity variable is not significant, whereas there is a strong evidence of positive self-selection in the earnings of non-migrants. This finding supports the notion that nonmigrants in the [rural] population choose their status because they fail to perceive more favorable returns elsewhere.

J.E.L. Classification : C25, F22.

Keywords : Human Capital, Probit Model, Rural-Urban Migration, Self-Selection, Tunisia.

[1] An earlier French version of this paper was presented at XIII Journées de micro-économie appliquée, Liège, Belgium, May, 1996.

* The authors wish to thank the anonymous reviewer and Michael Sibson for their comments and useful suggestions.

1-Introduction

For several decades, in developing countries, massive rural-urban migration has been observed[2]. There are several types of rural-urban migration : migration with or without a planned return, migration of single persons or of families, migration with or without a future work contract. Yet, in this paper,the differences between these types will not be considered. As a typical case, an individual rural-urban migrant without planned return and without a future work contract is assumed.

Despite more than 40 years of intensive research in the field of rural-urban migration, social scientists are still uncertain about the quantitative sources of the « explosive » city growth in the Less Developed Countries since the 1950s. The two principal hypotheses advanced in the literature are that rapid city growth and urbanization can be explained primarily by unusually rapid rates of population growth pressing on limited farm acreage, pushing landless labor into the cities; and economic forces pulling migrants into the cities.

Most demographers favor the first hypothesis. Exploding numbers of people must be employed, and a marginal agriculture sector with quasi-fixed arable land stocks cannot offer sufficient employment for a Malthusian glut created by the demographic urban transition.This view has also had a profound influence on economists' analysis thinking about development. It is central to Lewis's labor surplus model. It is also central to the Todaro thesis that rising immigration to the city is associated with high

---

[2] For a discussion of migration research regarding less developed countries, see Stark (1991), Stark and Bloom (1985) and Todaro (1976).

and even rising rates of urban unemployment[3]. On the other hand, most economists now tend to favor the second hypothesis; that is, an emphasis on those economic forces which contribute to urban pull. Thus, the motivating questions to be dealt with in the following are :

(1) What are the links between the « push » and « pull » causes of rural-urban migration in Tunisia; and what are the individual characteristics that distinguish migrants from the rest of the rural population ?

(2) Rationality dictates that persons choosing to migrate (the movers) do so because they have some tangible basis for perceiving a more favorable return than those who choose to stay (the stayers). The result is that movers tend to be non-randomly distributed within the rural population as a whole. As a consequence there is inherent « selectivity bias » which is a complicating factor in attempts to estimate returns to migration. Thus, what is the importance and what is the sign of the selectivity on unobservables in the rural-urban migration process in Tunisia ?

In the present paper, the individual rural-urban migration decision is modeled, focusing on the importance of human capital investment incentives and personal characteristics. An empirical test of the model is provided by adopting micro level data from rural areas of Tunisia. A test of the migration selectivity hypothesis is also undertaken : Do people who choose to migrate earn more in the urban locations than would identical individuals drawn at random from the rural population ? Positive selection will suggest that rural-urban migration decision is made according to the principal of comparative advantage.

---

[3] Bhattacharya (1993) provides a review of the theoretical literature on rural-urban migration in

The paper is organized as follows : Section 2 presents the Tunisian context. Section 3 describes a simple model of migrant behavior in which the decision to migrate, viewed in the context of investment in human capital, results as the solution to an optimal control problem. Section 4 provides details of a simultaneous equation model which incorporates the decision to migrate, returns to migration and self-selection. Data used and results of estimation are presented in Section 5. The paper concludes with a summary of the results.

2-The Tunisian Context

Between 1970 and 1992 Tunisia's population increased by 2.37 per cent per annum reaching almost 8.5 million inhabitants. Since 1972, and at each consecutive period, the labour force grew faster than the population indicating an increase in the female participation rate. Thus between 1972 and 1992 the labour force growth averaged 2.56 per cent, compared with a population growth of 2.4 per cent. This situation has of course aggravated the employment problem.

Table 1 : Population and labor force growth rates

Years	Population growth rate	Labour Force Growth Rate
1970-74	2.22 %	2.04 %
1975-79	2.64 %	3.06 %
1980-84	2.50 %	2.62 %
1985-89	2.36 %	2.14 %
1990-92	1.97 %	2.27 %

Source : Institut National des Statistiques, Tunis.

contemporary LDCs.

The year 1975 marks a turning point when the population was almost equally divided between rural and urban areas. Two decades previously, rural areas held twice as much of the population as the urban areas.

Rural-Urban migration accounted for the greater part of the shift in the distribution of the labour force (Radwan, Vali and Ghose, 1991). It has been one of the most important escape routes for the surplus agricultural labour. The present wave of migration, which dates from the 1970s, was associated with rapid mechanization of agricultural production and the development of the tourist industry in the coastel areas. These areas and the capital (Tunis) developed much faster than the northwest and the south. Rural labour from the disadvantaged areas was pushed to migrate to nearby urban centers, especially during the tourist season.

Table 2 : Labour Force Distribution Between Urban and Rural Areas

	1975	1984	1989
Labour Force in the Urban Areas (%)	51.11	54.59	62.28
Labour Force in the Rural Areas (%)	48.89	45.41	37.72
Unemployment Rate in the Urban Areas (%)	10.3	12.1	15.6
Unemployment Rate in the Rural Areas (%)	15.3	14.5	14.8

Source : Institut National des Statistiques, Tunis.

According to the 1984 population census, the inter-governorate migration amounted to 274 860 between 1979 and 1984. Intra-governorate migration amounted to 231 230 during the same period. The total (506 090) represented 7.3 per cent of the population of Tunisia or 23.7 per cent of the labour force. This points to a high degree of mobility of the Tunisian labour force which reflects in part the effect of regional differences.

Between 1984 and 1989 rural-urban migration represented almost 16 per cent of the total inter-governorate movement. During the same period, the rural-urban migration represented 9 per cent of the total intra-governorate movement.

3- An overview of the human capital migration model

The human capital model provides a way that, on an individual level, a decision is made to invest in human resources, in either additional education and training or the decision to migrate. The human capital framework suggests determinants of the decision that can be used in empirical analyses. Thus, the notion of migration as an investment in human capital remains a source of empirical interest almost thirty five years after Sjaastad's work (1962).

In this context migration is viewed as an investment through which income can be augmented. The basic concept weighs the benefits of moving against the cost of moving. In the simplest model, the benefits of migrating are represented by the present discounted value of the differences in lifetime earnings at the destination and at the origin. The individual's decision to migrate from rural *r* to the urban areas *u* depends on the calculation of the following objective function:

$$V_{ru}(t) = \int_{t_0}^{T} e^{-\rho(t-t_0)} (W_u(t) - W_r(t))dt - C_{ru}$$

where :

V_{ru} : net benefit of moving from rural areas *r* to urban areas *u*;

W_{rt} : expected rural income in period *t*;

W_{ut} : expected urban income in the destination area in period *t*;

ρ : subjective rate of time preference used as a discount rate;

C_{ru} : fixed costs of migration and relocation in the urban area;

t_0 : current time period (we suppose that $t_0 = 0$);

T : end of individual's working life;

t : the movement of time from the current time period (t_0) to T.

Thus, the basic hypothesis of this theory is that migration (from rural to urban areas) is largely in response to economic incentives. Following the discrete choice approach to migration, we can formally partition the (latent) net benefit of moving for individual i into two additive components:

$$M_i^* = V_{ru}^i + \varepsilon_i$$

where ε_1^i represents a portion of the variation in net benefit function which is unexplained and attributed to situational factors unknown to the researcher. Thus, the probability of migration P_i for the i individual is:

$$P_i = P(V_{ru}^i + \varepsilon_i > 0)$$

where

$$V_{ru}^i(t) = \int_0^{T_i} e^{-\rho_i t}(W_u^i(t) - W_r^i(t))dt - C_{ru}^i \text{ and } \frac{\partial P_i}{\partial V_{ru}^i} > 0.$$

If $W_u^i(t)$ and $W_r^i(t)$ are constant, the integral term in the net benefit's function becomes $(W_u^i - W_r^i)(1 - e^{-\rho T_i})/\rho$ and assuming that ρ is not large (non-myopic planning horizon) this expression can be approximated by $(W_u^i - W_r^i)T_i$. Introducing the cost component of migration, the probability of rural-urban migration P_i can be defined as a function of the wage differential, the time during which the individual will remain in the labor force (which is highly correlated with the age variable) and the costs variables which can vary with the observable characteristics of individuals and with regional attributes[4].

Refinements of the simple Sjaastad model include a model described by Polachek and Horvath (1977), who propose an optimal control model of life cycle locational change. The individual is assumed to maximize his present value of lifetime earnings, where available controls include : investment in human capital; a strategy of search for attractive wage opportunities in other locations; and mobility investment in the form of location change.

Polachek and Horvath (1977) have proposed an interesting refinements of the simple Sjaastad model. They have proposed an optimal control model of life cycle locational change. Under certain conditions : « the model gives rise to a conventional « bang-bang » solution, wherein the optimal control switches from its upper to lower bound without assuming intermediate values. The important feature of this dichotomy is that the decision to migrate may be modeled in part by means of a binary variable representing « move » or « stay » for each individual » (Nakosteen and Zimmer, 1980, p. 842).

The model presented hereunder, vigorously inspired by Nakosteen and Zimmer (1980), adheres to a similar convention, while incorporating endogenous selectivity into a model of rural-urban migration and income.

[4] I am grateful for the comment of anonymous referee concerning this derivation which has the advantage of linking the economic model and the econometric specification.

4-Econometric Model Specification

The model consists of two income equations (one for the mover, the other for the stayer) as well as an equation describing the dichotomous decision to migrate, where the dependent variable is the binomial observation on being in an urban area or not. The sample observations may be thought of as falling into one of two mutually exclusive regimes, with the decision equation serving as an endogenous selectivity criterion which determines the appropriate regime (rural-urban migrant versus non-rural-urban migrant or mover versus stayer). Thus, the model fall in the general class of switching models with endogenous switching (Maddala , 1983).

It is assumed that at any point in time, individual i decides to migrate if and only if the benefit in terms of the wage gap exceeds the cost. i.e., if and only if:

$$M_i^* = V_{ru}^i + \varepsilon_i > 0 \qquad (1)$$

where V_{ru}^i depends on wage gap, age and other individual and regional attributes defining the direct and indirect costs C_{ru}^i. These costs may be represented as a function of some personal characteristics X_i and locational factors Z_i .

Condition (1) suggests that the migrant selectivity criterion is a function of gains in earnings along with locational factors and personal attributes. This is expressed as[5]:

$$M_i^* = \alpha_0 + \alpha_1\left(Log(W_u^i) - Log(W_r^i)\right) + \alpha_2 AGE_i - \alpha_3 X_i - \alpha_4 Z_i - \varepsilon_i \qquad (2)$$

The model is completed by specifying income equations for movers and stayers respectively :

---

[5] If the relative change in income is relatively small, $Log(W_u^i) - Log(W_r^i)$ can approximate the percent gain in moving $(W_u^i - W_r^i) / W_r^i$.

$$Log(W_i^u) = \beta_{u0} + \beta_{u1} X_i^u + \beta_{u2} Z_i^u + \varepsilon_{ui} \quad (3)$$

$$Log(W_i^r) = \beta_{r0} + \beta_{r1} X_i^r + \beta_{r2} Z_i^r + \varepsilon_{ri} \quad (4)$$

where the disturbance terms ε_{ui} and ε_{ri} are assumed to be normally distributed with variances σ_{u}^2 and σ_{r}^2 respectively.

Equations (2)-(4) constitute the basic structural form of the model, where the endogenous variables are M_i^* and *Log(W_{ij}), j = u,r.* M_i^* is a latent variable, an index function, that represents the gains of migration. A dummy variable M_i takes the value 1 when $M_i^* > 0$ (the individual migrates), and takes the value of 0 otherwise. Given that agents can take only one of the decisions at a time, earnings for migrants are observed only when $M_i = 1$ and for non-migrants when $M_i = 0$.

The OLS is inappropriate for the income equations, since the conditional means of the income disturbance terms are non-zero and not constant for all observations. Indeed, using the well-known conditional formulae for the truncated normal distribution :

$$E(\varepsilon_{ui} / M_i = 1) = \sigma_{u\eta}[-\varphi(\lambda_i) / \Phi(\lambda_i)] \quad (5)$$

$$E(\varepsilon_{ri} / M_i = 0) = \sigma_{r\eta}[\varphi(\lambda_i) / 1 - \Phi(\lambda_i)] \quad (6)$$

where $\varphi(.)$ and $\Phi(.)$ are the standard normal density and distribution functions respectively, and $\sigma_{u\eta}$, $\sigma_{r\eta}$ and λ_i are defined below.

The argument λ_i is obtained as follows : substituting (3) and (4) into (2) giving the reduced form of the decision equation :

$$M_i^* = \gamma_0 + \gamma_1 X_i' + \alpha_4 Z_i' - \eta_i \equiv \lambda_i - \eta_i \quad (7)$$

where vectors *X'* and *Z'* consist of all exogenous variables (variable *AGE* included) in the model. Assuming that η is normally distributed with unit variance, (7) may be estimated by the maximum likelihood probit methods. Thus, probit estimation yields fitted values $\hat{\lambda}_i$, which are to be used as estimates of the arguments in (5) and (6). The coefficients $\sigma_{u\eta}$ and $\sigma_{r\eta}$ are elements of the covariance matrix of the disturbances:

$$\Sigma = \text{cov}(\varepsilon_u, \varepsilon_r, \eta) = \begin{pmatrix} \sigma_u^2 & \sigma_{ur} & \sigma_{u\eta} \\ & \sigma_r^2 & \sigma_{r\eta} \\ & & 1 \end{pmatrix}.$$

Note that the basic Roy (1951) model imposes restrictions on the signs of $\sigma_{u\eta}$ and $\sigma_{r\eta}$ in the wage equations (Maddala, 1983): Assuming no mobility costs or that unobersables in the wage equations are uncorrelated with those in the costs equation, the condition $\sigma_{r\eta} - \sigma_{u\eta} > 0$ must be satisfied. Assuming that this constraint is respected, we three cases can be considered concerning the sign of the covariances $\sigma_{u\eta}$ and $\sigma_{r\eta}$:

Case1 : $\sigma_{u\eta} < 0$ and $\sigma_{r\eta} > 0$. In this case those who choose to migrate are better (concerning the mean income of these persons) than average migrants, and those who chose to stay in the rural areas are are also better than the average stayers.

Case2 : $\sigma_{u\eta} < 0$ and $\sigma_{r\eta} < 0$. In this case those who chose to migrate are better than average in terms of mean income in both moving and staying, but they are better in moving than in staying in rural areas. Those who chose staying are below the average in both moving and staying, but they are better in staying than in moving.

Case3 : $\sigma_{u\eta} > 0$ and $\sigma_{r\eta} > 0$. This is the reverse of case 2.

The model described in this section explicitly recognizes the endogenous nature of the rural-urban migration decision and thus formally accounts for the problem of migrant self-selection.

5-Estimations Results

The data are drawn from a micro data set developed by the Tunisian averseas workers office (Office des Tunisiens à l'Etranger, OTE) in 1987. The sample survey elaborated provides general socioeconomic information on 10,000 households and approximately 50,000 individuals. More detailed information concerning either internal and international migration are reported on more than 4,000 individuals. Our analisys is confined to persons in the age group 15-55 in 1974 and who reported positive remuneration. Excluded from the sample are persons who have experienced international migration.

With respect to migration information, the survey reports the current (in 1987) and the last place of residence in 1974 in addition to the location where the person was born. It is also provides information on the years of residence in the present location.

Table 3 displays the definitions and measurements of the variables used in the analysis. The definition of rural-urban migration is inevitably a rather arbitrary one. We choose to define rural-urban migration as a movement from a rural area in the period 1974-1987 into an urban areas in the principal cities of Tunisia.

As noted above, the structural form of the model consists of a decision equation and earnings equations for migrants and non-migrants :

$$\begin{aligned} M_i^* &= \alpha_0 + \alpha_1\left(Log(W_u^i) - Log(W_r^i)\right) + \alpha_3 AGE_i + \alpha_4 AGE_i^2 + \alpha_5 MARIE_i \\ &+ \alpha_6 NSYEAR_i + \alpha_7 FEMAL * MARIE_i + \alpha_8 SACT0_i + \alpha_9 TYPRESD1_i + \alpha_{10} TYPRESD2_i \\ &+ \alpha_{11} TYPRESD4_i + \varepsilon_i \end{aligned}$$

$$Log(W_i^u) = \beta_{u0} + \beta_{u1} NSYEAR + \beta_{u2} DACTIVE1_i + \beta_{u3} IND87_i + \beta_{u4} TRAIN74_i + \varepsilon_{ui} ,$$

$$Log(W_i^r) = \beta_{r0} + \beta_{r1} EXP_i + \beta_{r2} EXP_i^2 + \beta_{r3} SECU87_i + \beta_{r4} FEMALE_i * MARIE_i + \varepsilon_{ri}$$

The variable *AGE* is included in the decision equation to reflect the widely held notion that the probability of migration declines with age. The concavity of the observed probability profile is captured by the quadratic age terms *AGE* and AGE^2 whose coefficients are expected to be respectively positive and negative. The coefficient of *FEMALE*MARIE* is expected to be negative to reflect the effect of family ties on the mobility of women[6]. *NSYEAR* may represent several effects of the education on the migration decision, including a higher probability of finding an urban job because of better access to urban labor market information and the use of educational attainment as a screening device by employers. Thus the coefficient of *NSYEAR* is expected to be positive.

Inclusion of all exogenous variables in both the decision and earnings equations introduces a collinearity problem in the second stage of the estimation procedure. In this case, those included are *EXP*, EXP^2, *TRAIN74, SECU87, IND87* and *DACTIVE1*. This last exogenous variable should have a negative impact on earnings, since individuals who change activity of employment are likely to experience depreciation in specific human capital periods immediately following the change.

---

[6] The influence of family ties on migration has been analyzed by Mincer (1978), who shows that, in the context of a nuclear family in which both husband and wife work, such ties tend to discourage migration. « Tied persons in the family are those whose gains from migration are dominated by gains of the spouse » [Mincer (1978,p.753)]. An important assumption of Mincer's model is that migration involves the movement of the entire family. This, however, would not be an accurate evaluation of migration in many LDCs where generally rural-urban migration does not involve the movement of the entire family but only of individual members who continue to retain strong bonds with the family members left behind in the rural sector [Bhattacharya (1993, p. 265)].

Substitution of the earnings equations into the decision equation results in a reduced form decision equation which includes as explanatory variables all the exogenous variables in the model. Maximum likelihood estimates of this equation are presented in Table 4.

Table 3 : Definitions of variables

Dependent variable in migration equation	
MIGRU2	= 1 if moved from rural to urban areas in the period 1974-87, 0 otherwise
Dependent variable in wage equations	
REV87	= Log of the real yearly earnings in Dinars.
Personal Characteristics	
AGE	= Age at time of migration for migrants and at the average age in period 1974-87 [(age in 74 + age in 87)/2] for non-migrants.
MARIE	= 1 if currently married, 0 otherwise.
FEMALE	= 1 if women, 0 otherwise.
Education and labor force	
NSYEAR	= Years of schooling.
TRAIN74	= 1 if did vocational training in 1974, 0 otherwise
SACT0	= 1 for those with an occupational skill not transferable to the urban economy, 0 otherwise.
DACTIVE1	= 1 if the individual changes activity sector's employment, 0 otherwise.
EXP	= Experience[7].
IND87	= 1 if the individual is self-employed in 1987, 0 otherwise.
SECU87	= 1 if the individual is employed in the public sector in 1987, 0 otherwise.
Locational variables	
TYPRESD1	= 1 if residence in 1974 was in Tunis, 0 otherwise.
TYPRESD2	= 1 if residence in 1974 was in the North, 0 otherwise.
TYPRESD3	= 1 if residence in 1974 was in the Center, 0 otherwise.
TYPRESD4	= 1 if residence in 1974 was in the South, 0 otherwise.

[7] In place of the standard experience variable (experience = age - nsyear - 6) and to avoid the problem of multicollinearity between *AGE*, *NSYEAR* and this standard experience variable, another *score* variable is included which also conveys the role of human capital. *EXP* is defined as the proportion of affirmative answers to the questions related to the individual vocational training in 1974 and 1987, and if this vocational training is in accordance with his present job.

Table 4 : Estimation of the reduced form

Variables	Coefficient	t-Student
Constant	-4.07	-6.13
AGE	0.17	4.11
AGE^2	-2.4E-03	-3.98
MARIE	-0.57	-3.42
NSYEAR	0.02	1.59
*DACTIVE*1	0.04	0.26
*IND*87	-4.7E-03	-0.03
EXP	0.28	0.25
EXP^2	-0.34	-0.23
*TRAIN*74	0.11	0.40
*SECU*87	0.20	1.35
*SACT*0	-0.58	-2.61
*FEMALE*MARIE*	0.33	1.54
*TYPRESD*1	0.53	3.10
*TYPRESD*2	0.39	2.36
*TYPRESD*4	0.35	2.18

Observation Number = 990
Migrants = 101 (10.2 %)
Non Migrants = 889 (89.3 %)
Log-likelihood = L(α) = -286
Log-likelihood at (α = **0**) = L(**0**) = -326
Khi-Squared (15) = 80

The Chi-Squared (13), equal to -2(L(**0**)-L(α)), is a statistic used to test the null hypothesis that all the parameters are zero. It is asymptotically distributed as χ^2 with 13 degrees of freedom. Here the value of the statistic is 80, which indicates that we can reject the null hypothesis that all the parameters are zero at the 0.01 level significance.

An informal goodness-of-fit index that measures the fraction of an initial log likelihood value explained by the model is defined as follows :

$$\bar{\rho}^2 = 1 - \frac{(L(\alpha) - K)}{L(0)} = 0.08,$$

this index is analogous to $\overline{R}^2$ used in regression, but it should be used with somewhat more caution (See Ben-Akiva and Learman, 1985).

Examination of the results reveals that the signs of parameter estimates generally conform to a priori expectations. The effect of *AGE* in the reduced form can be seen by using the marginal effects (slopes) and differentiating estimates of the reduced form with respect to *AGE* and setting equal to zero. This shows the maximum probability of migration occuring at age 36 years.

In the next step, fitted values from the reduced form probit model are used to construct selectivity variables, one for migrants and the other for non-migrants, in accordance with expressions (5) and (6). When these variables are appended to the corresponding earnings equations, the resulting corrected equations may be estimated by OLS. Estimates of the earnings equations are presented in Tables 5 and 6.

As noted in Nakosteen and Zimmer (1980, p.847-849), the combined effect of the two selectivity variables on unconditional earnings should be positive. Based on the estimates from Tables 5 and 6, we have $\hat{\sigma}_{r\eta} - \hat{\sigma}_{u\eta} = 1.70 - 1.12 = 0.58$, indicating that the combined effect on earnings is positive.

Of particular interest in this study is the sign and significance of the selectivity terms in the movers and the stayers group. The coefficients of the selection correction variables are significant at 10% and 1% significance level respectively. The t-values reported in Tables 5 and 6 are corrected to take into account the fact that the selection correction variables are estimated. The coefficients are positive (***Case 3***) : stayers earned more, ceteris paribus, in the origin than the movers would have done if they had stayed. Thus, this is evidence of positive selection in the two earnings equations and especially for non-migrants. Thus, non-migrants in the population choose their status because they fail to perceive more favorable returns elsewhere.

Table 5 : Migrant Earnings Equation[8]

Variables	Coefficients	T-Student
Constant	5.80	4.78
NSYEAR	0.06	1.93
DACTIVE1	-0.83	-1.77
IND87	0.73	1.28
TRAIN74	0.17	0.34
$-\varphi(\hat{\lambda}_i) / \Phi(\hat{\lambda}_i)$	1.12	1.67
Observations Number = 101		
R^2 = 0.11		
Log-likelihood = -201		

[8] As noted by Borjas (1989, p.467): It is well known that the explanatory power of earnings functions [like equation (3) and (4)] is relatively small, regardless of the specification of the vector [of regressors]. It is not uncommun, for example, for an earnings function of this type to have an R^2 of less than 20 or 30 percent. ... Unobserved differences among individuals either in skills, ability or luck dominate earnings distribution data.The low quality of the regression concerning migrant earnings equation can also be explained by the limited number of observations (101).

Table 6 : Non-Migrant Earnings Equation

Variables	Coefficients	T-Student
Constant	1.47	6.80
EXP	7.73	7.09
EXP^2	-6.15	-5.00
SECU87	0.99	6.15
*FEMALE*MARIE*	-1.32	-5.87
$\varphi(\hat{\lambda}_i) / (1-\Phi(\hat{\lambda}_i))$	1.70	3.34
Observations Number = 889		
R^2 = 0.25		
Log-likelihood = -1768		

The final step in the Heckman two-step estimation procedure (Heckman, 1979) entails probit estimation of the structural form of the migrant status equation. Consistent estimates of the parameters in the earnings equations are used to obtain fitted values of log-earnings, which together with appropriate exogenous variables are then inserted into the structural decision equation. The resulting maximum likelihood estimates are presented in Table 7.

The second point of interest concerns the estimated coefficient of the earnings variable $\left(Log(\hat{W}_u^i) - Log(\hat{W}_r^i)\right)$. Its positive and significant value lends strong support to the essential hypotheses of the conventional human capital model of migration.

The structural estimated parameters of the model, like those of any nonlinear regression model, are not necessarily the marginal effects. Indeed,

$$\begin{aligned}\frac{\partial E(M)}{\partial X} &= \left\{\frac{\partial \Phi(\beta' X)}{\partial(\beta' X)}\right\}\beta \\ &= \varphi(\beta' X)\beta\end{aligned}$$

where $E(M) = 0[1-\Phi(\beta' X)] + 1[\Phi(\beta' X)] = \Phi(\beta' X) = P(M=1)$, β is the set of unknown parameters in the strucutral equation, X is a vector of individual-specific characteristics, φ*(.)* is the standard normal density function that corresponds to the cumulative distribution, Φ*(.)*.

It is obvious that these will vary with the values of *X*. Thus, in interpreting the estimated model, it will be useful to calculate using the means of the regressors. The resulting marginal effects computed at the means of the *X*s are presented in table 8.

We shall shortly comment on the statistical and economic features of the estimation results concerning the earnings variable which is continuous. The effect of $(Log(\hat{W}_i^u) - Log(\hat{W}_i^r))$ is statistically very significant but its influence alone on rural-urban migration is very small (5.2%) as illustrated in **Figure 1**.

The coefficient of earnings variable can be severely downward biased by the omission of other relevant variables from the analysis such as entry into the labor force, start of a career, birth and aging of children, characteristics of receiving urban regions (job opportunities, unemployment rate...), repeat and return migration experience. But these variables were not available in the data to which we had access.

Table 7 : Probit Estimation of the Structural Form

Variables	Coefficient	t-Student
Constant	-6.16	-8.80
AGE	0.22	5.08
AGE^2	-3E-03	-4.83
MARIE	-0.61	-3.52
NSYEAR	0.04	3.59
*FEMALE*MARIE*	-0.50	-2.03
SACT0	-0.92	-3.88
TYPRESD1	0.53	2.91
TYPRESD2	0.44	2.39
TYPRESD4	0.44	2.53
$(Ln(\hat{W}_i^u) - Ln(\hat{W}_i^r))$	0.53	8.26

Observation Number = 990
Migrants = 101
Non Migrants = 889
Log-likelihood = $L(\alpha)$ = -249
Log-likelihood at ($\alpha = 0$) = L(0) = -326
Khi-Squared (15) = 155
$\overline{\rho}^2$ = 0.21
% of good predictions = 89.5

Table 8 : Marginal Effects

Variables	Coefficient	t-Student[1]
AGE	0.022	4.973
AGE^2	-3E-04	-4.811
MARIE	-0.059	-3.420
NSYEAR	0.004	3.457
*FEMALE*MARIE*	-0.049	-1.995
SACT0	-0.091	-4.225
TYPRESD1	0.051	2.888
TYPRESD2	0.042	2.433
TYPRESD4	0.043	2.565
$(Log(\hat{W}_i^u) - Log(\hat{W}_i^r))$	0.052	6.812

[1] Marginal effects and the asymptotic covariance matrix are computed at the means of the regressors [See Greene (1993, Chap. 21)]

Erreur! Liaison incorrecte.

Figure 1 Fitted probability of rural-urban migration as a function of earnings variable

6. Conclusions

The chief purpose of this paper was to model and test individual rural-urban migration decision in Tunisia and the migration selectivity, focusing on the importance of human capital investment incentives and personal characteristics.

The results, using a micro level data from rural areas of Tunisia, lead to the conclusion that by purely statistical assessment the expected monetary gains effect is significantly different from zero, but that by economic considerations it is small. This low effect can be explained by the omission of other relevant variables from the analysis of rural-urban migration in Tunisia.

In the migrant-earnings equation the selectivity variable is not significant, whereas there is a strong evidence of positive self-selection in the earnings of non-migrants. As in Nakosteen and Zimmer (1980, p.847), this finding supports the « notion that nonmigrants in the [rural] population choose their status because they fail to perceive more favorable returns elsewhere ».

REFERENCES

Bhattacharya, P. C., 1993 : « Rural-urban Migration in Economic Development », *Journal of Economic Surveys*, vol 7, n°3, p. 243-281.

Ben-Akiva, M. and S. R. Lerman : « Discrete Choice Analysis : Theory and Application to Travel Demand ». The MIT Press, Cambridge, Massachusetts, London.

Borjas, G. J., 1989 : « Economic Theory And International Migration », *International Migration Review,* Vol. xxiii, n°3, p. 457-485.

Greene, W. H., 1993 : *Econometric Analysis*, 2nd Edition, Prentice-Hall, Inc.

Heckman, J.J., 1979 : « Sample Selection Bias as a Specification Error », *Econometrica*, 47, p.153-161.

Maddala, G. S., 1983 : « Limited Dependent and Qualitative Variables in Econometrics». Cambridge University Press.

Nakosteen, R.A. and M. Zimmer, 1980 : « Migration and Income : The Question of Self-Selection », *Southern Economic Journal*, 46 (3), p.840-851.

Polachek,S.W. and F.W. Horvath, 1977 : « A Life Cycle Approach to Migration : Analysis of The Perspicacious Perigrinator ». in Research in Labor Economics, Ehrenberg, Greenwich editors : JAI Press.

Radwan S., Vali J. and A. Ghose, 1991 : « Tunisia : Rural Labour and Structural Transformation ». Routledge Ed. International Labor Organisation.

Roy, A. D. (1951) : « Some Thoughts on The Distribution of Earnings », *Oxford Economic Papers*, 3, p.135-146.

Sjaastad,L., 1962 : « The Costs and Returns of Human Migration. », *Journal of Political Economy*, October, p.80-93.

Stark, O. 1991, : The Migration of Labour. Oxford and Cambridge, MA. Basil Blackwell.

Stark, O. and D. E. Bloom, 1985 : « The New Economics of Labour Migration, » *American Economic Review*, 75, p.173-178.

Todaro, M.P., 1976 : Internal Migration in Developing Countries : A Review of Theory, Evidence, Methodology and Research Priorities. Geneva, International Labour Office.