INTERNATIONAL POLICY CENTER
Gerald R. Ford School of Public Policy
University of Michigan

IPC Working Paper Series Number 41

# Migration, Self-selection and Returns to Education in the WAEMU

**Philippe De Vreyer**
**Flore Gubert**
**Francois ROuband**

May 15, 2007

# Migration, Self-selection and Returns to Education in the WAEMU

Philippe De Vreyer*, Flore Gubert† and François Roubaud‡

Work in progress

May 15, 2007

**Abstract**

Using data from labour force surveys conducted simultaneously in the capital cities of seven WAEMU countries, we estimate a model of residential location choice, in which expected earnings play a role. The model is first estimated in a reduced form. Estimates are then used to correct for the endogeneity of locational choice in earnings equations estimated for each country. We find that migration behaviour has a significant effect in shaping earnings differentials between education levels and between the seven capital cities. A minimum distance estimator is then used to recover the value of log-earnings in the structural model of residential location choice. Results show that individuals tend to reside in countries in which their expected earnings are higher than elsewhere.

*Corresponding author. University of Lille II and DIAL, Paris. email: devreyer@dial.prd.fr
†IRD, DIAL, Paris.
‡IRD, DIAL, Paris.

# 1 Introduction

Migration from and to African countries is an extensive phenomenon. According to recent estimates by the United Nations Population Division, the total number of international migrants in Africa rose from nine millions in 1960 to 16 millions in 2000. West Africa in particular has a long history of population mobility, both regionally and internationally. Linked with factors as diverse as long-distance trade, plantation agriculture, urbanisation but also armed conflict, land degradation, drought, etc., migration in the region played and still plays a major part in shaping settlement patterns. At a political level, several initiatives have facilitated labor migration, among which the free movement of persons institutionalized by the Economic Community of West African States (ECOWAS).

With this background in mind, the purpose of this paper is to examine the locational choice of a large sample of Africans originating from the West African Economic and Monetary Union (WAEMU). Historically, in the economic literature the concern with migration emerged with the work of Sjaastad (1962). In the development literature, however, Todaro (1969) and Harris & Todaro (1970) are the first to present a model in which the decision to migrate results from the rational comparison of the expected costs and benefits of migration. In both models, the difference in average expected earnings between countries or regions of destination and countries or regions of origin plays a key role and is predicted to have a positive effect on migration flows. However this kind of model is unable to explain key stylized facts, such as migration flows from and to particular regions or countries. For instance, in Africa, a sizable number of people living in Benin come from Togo and an equally sizable number

of people residing in Togo are natives from Benin. Borjas (1987) and, more recently, Dahl (2002) have adopted a rather different approach, based on the seminal paper of Roy (1951). In Roy's framework, workers select themselves in income earning activities on the basis of their comparative advantage. Applied to residential choice, this model explains migration not by average expected earning differentials, but rather by differences in individual expected returns to skills that are either observed or unobserved by the econometrician. As a result migration flows are not necessarily one-sided. Another conclusion of this literature is that migrants' self selection should be taken into account when estimating the returns to human capital in countries where the flow of migrants is significant. Dahl (2002) for instance, in a study of migration between states of the USA, estimates a Roy model and finds that correcting for selection bias substantially changes the estimated returns to education in a sense that supports the role of comparative advantage in mobility decisions. He also finds that migration flows depend positively on the differences in the corrected returns to education.

Estimation of this kind of model is usually very difficult due to the impossibility to gather data on the origin and destination labour markets at the same time. In this paper we use a unique collection of data originating from the PARSTAT project sponsored by the WAEMU.[1] Representative household quantitative surveys have been conducted simultaneously in the capital cities of seven member States of the WAEMU (Abidjan, Bamako, Cotonou, Dakar, Lome, Ouagadougou and Niamey) in 2001-2002. The surveys provide detailed

[1] The PARSTAT project was coordinated by AFRISTAT, under the scientific supervision of Alain Brilleau (DIAL-INSEE), Eloi Ouedraogo (AFRISTAT) and François Roubaud (DIAL-IRD). See Amegashie, Brilleau, Coulibaly, Koriko, Ouedraogo, Roubaud & Torelli (2005) for details on the project and Brilleau, Roubaud & Torelli (2005) for extensive descriptive results.

information for all individuals aged 10 or more within each sample household, relating to education and training, employment, unemployment and earnings. Furthermore, data on country of birth and last country of residence allow to identify international migrants within each national sample.

Our purpose in the paper is threefold. First, we fill a gap in the knowledge of cross-border migrations within Africa, using our sample data to compare the characteristics of migrants with those of non migrants in their countries of origin and destination. Second, we want to evaluate the extent of the bias in the estimated returns to education, when international migration is not accounted for. Third, we want to determine whether or not earnings differentials matter in the choice of the country of residence. In the model that follows we assume that individuals are born randomly in one of the seven countries under review, but then rationally choose the country in which they reside by comparing the utilities associated with each choice. Estimation of this model provides unbiased estimates of the returns to education, together with the effect of expected earnings differentials on the probability of choosing one particular country. We find that migration behaviour plays an important role in determining earnings differentials between countries and between individuals with different education levels. Moreover, our results suggest that earnings differentials matter in locational choice.

## 2 Data and descriptive statistics

Movements of labour in Sub-Saharan Africa are not a new phenomenon. Over the generations people have migrated in response to demographic, economic,

political an other related factors, such as population pressure, environmental disasters, poverty and conflicts. In pre-colonial West Africa, migrations were generally circular, seasonal and of short duration, and occurred largely from unsecure or drought-prone regions to more secure and fertile regions (Adepoju 2005). Colonialism significantly altered the motivation and migration patterns in this region by introducing far reaching structural changes. In particular, the development of transportation systems, the monetization of the economy and the deliberate development of mining enclaves and plantation agriculture together with a series of recruitment policies (compulsory recruitment, contract and forced labour legislation and agreements) stimulated regional labour migration from Mali, Togo and Upper Volta to Gold Coast and Côte d'Ivoire (Adepoju 2005, Adebusoye 2006). These socio-economic and historical factors have shaped contemporary patterns of migration between African countries. However, with the end of colonialism and largely in response to growing disparities in living standards, inter-continental migration in the direction of Northern developed countries has been a growing phenomenon for the last fourty years.

Despite their importance, yet little is known about these migrations. The information provided by census data, immigration and emigration statistics and a small number of *ad hoc* surveys on the number, identity and motivations of both inter- and intra-continental African migrants is indeed far from being complete and reliable. In particular, estimates on the number of African international migrants widely differ between sources: they range from about 16 million according to the International Organisation for Migration (IOM, 2003) to 50 million according to the African Union (AU, 2005). Evidence is even more scarce concerning trans-border migrations within the West African sub-region.

How many trans-border migrants are there in each West-African country? Who are these migrants? What are their main motivations ? Here are some of the questions we want to address in this paper.

Our data come from representative household quantitative surveys (the 1-2-3 Surveys on Employment, Informal Sector, Consumption and Poverty) conducted simultaneously in the capital cities of Benin (Cotonou), Burkina Faso (Ouagadougou), Cote d'Ivoire (Abidjan), Mali (Bamako), Niger (Niamey), Senegal (Dakar) and Togo (Lome) in 2001-2002. These countries are all members of the Economic Community of West African States (ECOWAS).[2] The creation of ECOWAS, in 1975, responded to the recognition by West African leaders that intra-regional integration could be an important step towards the region's collective integration into the global economy. The key objective of the Community was thus to remove obstacles to the free movement of goods, capital and people in the sub-region. In line with this objective, the Protocol on Free Movement of Persons and the Right of Residence and Establishment was signed in May 1979. A transition period followed, during which the rights of entry (in 1980) and residence (in 1986) were established. More recently, in 2000, members of the ECOWAS agreed to introduce a new passport for citizens of the sub-region that will progressively replace national passports. Even though much remains to be done in order to achieve a complete liberalization of labour migration within the community - some countries are still restricting foreigners, including community nationals, from participating in certain kinds of economic activities - all these measures taken to create a borderless West Africa provide a good

[2]ECOWAS groups 15 countries: 5 English speaking countries (Gambia, Ghana, Liberia, Nigeria, Sierra Leone), 8 French speaking countries (Benin, Burkina Faso, Guinea, Ivory Coast, Malin, Niger, Senegal, Togo) and 2 sharing Portuguese as their official language (Guinea Bissau and Cape Verde).

opportunity to study the residential choice of people within the community. Moreover, amongst the ECOWAS members, the countries of our sample are all French-speaking countries and are all members of another community, namely the West African Economic and Monetary Union (WAEMU).[3] As such they share the CFA franc as a common currency. These common features undoubtedly facilitate labour migration.

Implemented by National Statistical Institutes in conjunction with AFRISTAT and the IRD Research Unit DIAL, the 1-2-3 Surveys provide detailed information for all individuals aged 10 or more within each sample household relating to education and training, employment, unemployment and earnings. Furthermore, data on country of birth and last country of residence allow us to identify migrants within each national sample. More details on the survey can be found in Amegashie et al. (2005).

Table 1 reports the composition of each national sample. For ease of computation, are considered as natives of country $i$ all individuals who have resided in country $i$ on a permanent basis, whether they declare having country $i$'s citizenship or not. In the empirical analysis that follows, we restrict the sample to all active individuals aged 16 or more, originating from one of the seven countries covered by the 1-2-3 survey and residing in the capital city of one of these countries either as natives or as immigrants. To avoid confusion, all indi-

[3]Created in 1994, the West African Economic and Monetary Union (WAEMU) is composed of eight member States: Benin, Burkina Faso, Côte d'Ivoire, Guinea-Bissau, Mali, Niger, Senegal and Togo. Some of the principal objectives of WAEMU are to: (i) strengthen competitiveness of the economic and financial activities of the member States within the context of a free and competitive common market and a rationalised and harmonised legal environment; (ii) achieve convergence of the performance and economic policies of the member countries; and (iii) create a common market among the member countries based on free movement of persons, goods, services, and capital and the right of establishment of persons engaged in an independent or salaried employment, and on a common external tariff and trade policy.

viduals included in the sample appear in bold in Table 1. As suggested by the figures, there is a wide variety of migration configurations within the WAEMU. Figures first suggest that despite the severe sociopolitical crisis that started in 1999 with a military coup d'Etat, Cote d'Ivoire is still, by far, the most important immigration country in the WAEMU region.[4] Extrapolation from the Ivorian sample reveals that 15.9 per cent of Abidjan's inhabitants aged 16 or more are immigrants among which 74 per cent are citizens of a WAEMU country (see Table 2 for extrapolated figures). Even though migration flows from Burkina Faso and Mali have been fluctuating since the beginning of the crisis, these two neighboring countries remain the main providers of migrants to Cote d'Ivoire. By contrast, immigrants from bordering countries only account for a marginal share of the population in Ouagadougou, Bamako and Dakar. Last, a quick comparison of row and column totals by country suggests that Malian and Burkinabe expatriates residing in the capital city of a WAEMU country largely outnumber the expatriates from WAEMU countries residing in Bamako or Ouagadougou, suggesting that Mali and Burkina Faso have been and still are major labour-exporting countries. Benin and Togo, by contrast, combine both emigration and immigration.

Table 3 provides some descriptive statistics on the main characteristics of natives and immigrants by country of residence. Figures first suggest that compared to natives, females are under-represented in the immigrant population of Ouagadougou, Abidjan and Lome while they are slightly over-represented in that of Cotonou, Bamako and Niamey. Traditional male-dominated short-to-long distance migratory streams in West Africa are thus increasingly feminised,

[4] The civil war in Côte d'Ivoire started in september 2002, a few months after the completion of the 1-2-3 survey.

suggesting a turn-around in traditional sex roles. Second, no clear pattern emerges with regard to age. Immigrants are significantly older on average than natives in Abidjan and Niamey but are roughly of the same age in all the other capital cities. Third, immigrants appear to be less educated on average than natives in four capital cities out of seven (Cotonou, Abidjan, Niamey and Lome). As a result, the percentage of non graduate individuals among migrants in these four cities is much higher than among natives. The education gap is particularly pronounced in Abidjan where immigrants have two years of schooling on average against 6.6 for natives. By contrast, with more than eight years of schooling, immigrants in Dakar appear much more educated on average than natives. Due to small sample size, however, this last figure should to be taken with caution.

As a complement to Table 3, Table 4 provides some descriptive statistics on the main characteristics of non-migrant natives and emigrants, or "stayers" and "movers", by WAEMU country. In most countries, males are over-represented in the emigrant population except in Togo and, to a lesser extent, in Benin. Intra-regional migratory flows from these two countries are mostly motivated by commercial purposes and have traditionally been female-dominated. In terms of education, emigrants appear much less educated than non-migrant natives in all countries, suggesting that migration flows within the WAEMU region mainly concern low-qualified workers.

Last, Table 5 provides descriptive statistics on the employment situation of natives and immigrants by country of residence. On average, labour force participation is higher for immigrants than for natives. The difference is particularly strong in the cases of Abidjan and Niamey, suggesting that migration streams to these two capital cities are mainly motivated by labour market considerations.

Given the individual characteristics of immigrants, particularly with respect to their level of education, one would expect their employment situation to be less favourable than that of natives in Cotonou, Abidjan, Niamey and Lome and more favourable in Dakar. In the context of labour markets in developing economies, a favourable situation is that of formal wage workers in the public or private sector, in contrast to the situation of informal workers. Formal wage workers usually enjoy higher wages, more job security and more benefits than informal workers. Figures indicate that this is indeed the case. The percentage of immigrants working in the informal sector is much higher than that of natives in Cotonou, Abidjan, Bamako, Niamey and Lome while it is lower in Dakar and in Ouagadougou. Average hourly earnings roughly follow the same pattern. Compared to natives, immigrants are indeed found to enjoy much lower hourly wages on average in Cotonou (-26%), Abidjan (-36%) and Niamey (-29%) while they enjoy much higher hourly wages in Dakar (+172%), Lome (+39%) and Bamako (+28%). Figures for Dakar and Bamako should however be considered with great care given small sample size. Lome stands as an exception since its immigrants are less educated on average, are more concentrated in the informal sector but enjoy significantly higher hourly wages than natives.

# 3 Model specification and estimation strategy

We study the locational choice of individuals originating from one of the seven countries of the PARSTAT project. Each individual has the choice to settle in any of these seven countries. We assume that individuals behave as if they maximize a stochastic utility function, where utility is a function of the distri-

bution of earnings in the chosen location. The question is whether differences in individual specific mean earnings determine locational choice. The difficulty is that, since we observe earnings at only one location for each individual, potential earnings at other locations must be imputed and, in doing so, it is necessary to hold account of the fact that location choice is not random, but partly commanded by earnings differences. Thus our estimation strategy proceeds in three steps. In the first step a multinomial logit model of locational choice is estimated using a reduced form specification. The results from this estimation are then used to compute appropriate correction terms that are added as independent variables in Mincer-type earnings equations. Results from this second step are then used to identify the effect of expected earnings differentials in locational choice.

We assume that individual $i$, born in country $j$, and living in country $k$ has a utility $u_i(j,k)$ given by :

$$u_i(j,k) = \alpha . \ln y_{ik} + z_i' \gamma_k + v_i(j,k) \tag{1}$$

with $\ln y_{ik}$ the logarithm of the individual's hourly earnings in country $k$ and $z_i$ a vector of individual characteristics. An increase in labour market earnings provides identical gains in utility, independently of the country of residence. This might be too strong an assumption if large differences exist between countries in the set of available goods and their price. For instance health services could be free of charge in one country and very costly in another. This would impact on the living standards of people with identical incomes but not living in the same country. In the present case, the data we use come from very similar countries: all of them are former French colonies and they share a common cur-

rency. Moreover, all surveyed individuals live in capital cities, between which differences in markets are likely to be smaller than between urban and rural areas. In addition to earnings, we assume that utilities are impacted by individual characteristics, $z_i$, with the size and sign of the impact depending upon the country of residence. For instance, countries in the WAEMU largely differ by their population's religious composition: more than 90% of the population living in Dakar (Senegal), Bamako (Mali) and Niamey (Niger) is muslim, against about 10% in Lome (Togo) or Cotonou (Benin). *Ceteris paribus*, individuals of a given confession might prefer to live in countries where this confession is well represented. As a result, being a Muslim should have a positive impact on utility for people living in Dakar, Bamako and Niamey, but a zero or even a negative impact for people living in Lome or Cotonou.

Individual $i$ decides to live in country $k$ if this choice provides more utility than living in any other country, that is:

$$u_i(j,k) \geq u_i(j,l) \qquad \text{for any } l. \tag{2}$$

We are particularly interested in estimating $\alpha$ in equation (1). Since $\ln y_{ik}$ is only observed for individuals living in country $k$, estimation has to proceed in several steps. First, we assume that each individual living in country $k$ faces a Mincer-type earnings equation:

$$\ln y_{ik} = x'_{ik}.\beta_k + u_{ik} \tag{3}$$

where $x_{ik}$ is a vector of individual characteristics such as sex, education or labour market potential experience. Second, we substitute $\ln y_{ik}$ in equation (1)

and get utility in a reduced form:

$$u_i(j,k) = \alpha.(x'_{ik}.\beta_k) + z'_i\gamma_k + \varepsilon_i(j,k)$$

where $\varepsilon_i(j,k) = \alpha u_{ik} + v_i(j,k)$.

Under the assumption that $\varepsilon_i(j,k)$ has a generalized extreme value distribution, it can be shown that:

$$\begin{aligned} P(i \text{ lives in } k) &= P(M_i(j) = k) \\ &= \frac{\exp(\alpha.(x'_{ik}.\beta_k) + z'_i\gamma_k)}{\sum_{l=1}^{P} \exp(\alpha.(x'_{il}.\beta_l) + z'_i\gamma_l)} \\ &= \frac{\exp(x'_{ik}.\beta_k^{\alpha} + z'_i\gamma_k)}{\sum_{l=1}^{P} \exp(x'_{il}.\beta_l^{\alpha} + z'_i\gamma_l)} \end{aligned} \tag{4}$$

with $P$ the total number of locations and $\beta_k^{\alpha} = \alpha\beta_k$. This is known as the multinomial logit model and is well documented in standard reference textbooks.

Results from this reduced form estimation can then be used to correct for endogenous selection in the earnings equations.[5] As shown by Lee (1983), it is possible to adapt the two steps method suggested by Heckman (1979) to the case of polychotomous choice models. His intuition is that the dimension of the problem can be reduced by substituting the $P$ selection equations in (2) by the single condition that:

[5] The multinomial logit suffers from the Independance of Irrelevant Alternatives assumption, which in most cases is unlikely to hold. However based on Monte-Carlo simulations, Bourguignon et al (2004) conclude that *"selection bias correction based on the multinomial logit model seems a reasonable alternative to multinomial models when the focus is on estimating an outcome over selected populations rather than on estimating the selection process itself. This seems even true when the IIA hypothesis is severely at odds."* We are then confident that our choice of the multinomial logit should not bias our results.

$$\max_l(u_i(j,l) - u_i(j,k)) \leq 0$$

Then, transforming to normal the cumulative distribution function of the maximum order statistic achieves the transformation of the P-dimensional joint distribution of the earnings and selection equations error terms to one of a bivariate normal distribution, in which the Heckman procedure can be applied. However, as shown by Schmertmann (1994) and more recently by Dahl (2002) and Bourguignon, Fournier & Gurgand (2004), Lee's method implies very strong restrictions on the correlation structure of the earnings and selection equations disturbances. Dahl (2002) suggests a non parametric method that is less demanding. Moreover, as shown by Bourguignon et al. (2004), Lee's method is only adapted to very small samples so that non parametric methods should be preferred for larger samples . In this paper we thus use Dahl's correction method and Bourguignon et al. (2004)'s Stata program to estimate our model.

In a last step, we compute a minimum distance estimator to recover the value of the coefficient of log-earnings in the structural model of residential choice (Gourieroux & Monfort 1995). According to the structural model, the following set of constraints has to be satisfied between coefficients of equations (1), (3) and (4):

$$\widehat{\beta_k^\alpha} - \alpha.(\widehat{\beta}_k - \widehat{\beta}_0) = 0 \text{ for } k = 1 \text{ to } 6. \tag{5}$$

where index $'0'$ refers to the reference country in the reduced form multinomial logit equation.[6] This is a system of $6K$ equations, where $K$ is the number

[6] In the multinomial logit model, only the differences $\beta_k^\alpha - \beta_0^\alpha$, where 0 is the index of a

of variables in the $x_{ik}$ vector. Let $\theta = (\widehat{\beta_1^{\alpha}}', .., \widehat{\beta_6^{\alpha}}', \widehat{\beta}_0', .., \widehat{\beta}_6')'$ be the vector of estimated coefficients in the first and second steps of the estimation. The constraints system of equation can be written:

$$g(\widehat{\theta}, \alpha) = 0$$

Following Gourieroux & Monfort (1995), we can estimate $\alpha$ using asymptotic least squares. The optimal value of $\alpha$, $\widehat{\alpha}$, verifies:

$$\widehat{\alpha} = \arg\min(g(\widehat{\theta}, \widehat{\alpha})'.S_n.g(\widehat{\theta}, \widehat{\alpha}))$$

with $S_n = \left[\frac{\partial g(\widehat{\theta},\widehat{\alpha})}{\partial \widehat{\theta}'}.\widehat{\Omega}.\frac{\partial g(\widehat{\theta},\widehat{\alpha})'}{\partial \widehat{\theta}}\right]^{-1}$ and $\widehat{\Omega} = Var(\widehat{\theta})$.[7] In order to account for the fact that the estimation is done in several steps, the entire procedure has been bootstrapped with 50 replications and bootstrapped standard errors used whenever necessary.

# 4 Model identification and choice of variables

In order to be identified, our model relies on various assumptions that need to be properly tested. In particular, in the second step of our procedure in which we correct for individuals' self-selection, it is important to have one or more variables that explain locational choice (i.e. that enter the first stage equation) but do not influence earnings. In what follows, we use dummies for the individual's religion and nationality as identifying variables. Religion is indeed likely to have an influence on destination choice given that large differences

reference country, can be identified.

[7] The variance of the estimator is given by: $V(\widehat{\alpha}) = \left[\frac{\partial g(\widehat{\theta},\widehat{\alpha})'}{\partial \widehat{\alpha}}.S_n.\frac{\partial g(\widehat{\theta},\widehat{\alpha})}{\partial \widehat{\alpha}'}\right]^{-1}$.

exist between countries in their population's dominant religions. Nationality dummies are also included to account for macro-level variables, such as average GDP per capita, mortality rates or the shares of immigrants from ECOWAS countries in the country's population. We test these exclusion restrictions by including in turn religion and nationality dummies in the list of explanatory variables in the earnings equations. The joint significance of excluded regressors is then tested by a Wald test. When both sets of regressors are added, the model remains identifiable, however identification can be weak because it rests on the non-linearity of the correction terms in the earnings equations.[8] Nonetheless, we use this property to test for the validity of our exclusion restrictions.

In the third stage of our procedure, identification of the log-earnings coefficient, $\alpha$, in the structural model of residential choice depends upon the exclusion from equation (1) of at least one variable that enters in the log-earnings equation (3). Here we assume that sex, education and employment sector explain log-hourly earnings but not residential choice, once earnings are accounted for. There are some good reasons for which education could determine residential choice, apart from its impact on potential earnings. One possibility is that well educated individuals might prefer countries where the average level of education is high, not only because their own wages are going to be higher, but also because they will benefit from positive externalities related to this high average level of education (such as a higher supply of cultural goods for instance). In our case, however, since the average level of education is low in all the capital cities of our sample, we believe such incentives to be small.

In the earnings equation our dependent variable is the logarithm of total

[8] See, for instance, Wooldridge (2002)

hourly earnings in CFA francs. All earnings are expressed in purchasing power parity (PPP). The conversion to PPP CFA francs is necessary in the third step of our estimation, where the individual's expected earnings in the seven countries are allowed to influence the probability of choice. The PPP conversion factors we use are those published by the World Bank in its World Bank Indicators (World Bank 2003). Independent variables in the earnings equations are sex, education (as measured by the last diploma obtained), potential labour market experience and its square, the abilities to speak french and another foreign language, two dummies for the public or private formal sectors and a series of dummies for the father's activity when the individual was 15. This last set of variables is included both as a determinant of migration behaviour and as a proxy for the individual's sector choice, to account for the earnings differentials between the different sectors of the economy. The reduced form multinomial logit model includes these variables, together with dummies for the individual's religion and nationality.

# 5 Estimation Results

## 5.1 Reduced form multinomial logit of residential choice

Estimation results are presented in Tables 6 and 7. Table 6 shows the results of the reduced form multinomial logit estimation. These are uneasy to comment because only the differences $\beta_k^\alpha - \beta_0^\alpha$ can be identified, where 0 is the index of a reference country (Senegal in our case). Thus, for instance, the positive coefficient of the sex variable in the equation for Benin tells that being a male increases *relatively* more the utility resulting from choosing Benin than the util-

ity resulting from choosing Senegal. However it does not mean that being a male increases the utility associated with Benin in absolute terms, as it could happen that $0 > \beta_k^\alpha > \beta_0^\alpha$. The results suggest that, among the seven countries under review, holding the baccalaureate increases more (or decreases less) the utility to reside in any other country than that of residing in Senegal. By contrast, holding a postgraduate degree increases the utility of residing in Senegal much more than that of residing in any other country. The same holds true for people of muslim or catholic confession. Unsurprisingly, we also find that being of Senegalese nationality increases much more the utility to reside in Senegal than that of residing in any other country, with the exception of Mali, but the coefficient is insignificant (results not shown).

## 5.2 Earnings equations

Following Dahl (2002), the estimated coefficients of the reduced form multinomial logit have been used to compute, for each observation of the sample, a vector of choice probabilities. A two-order polynomial of these probabilities has then been added to the list of explanatory variables in the Mincer-type earnings equations in order to correct for the endogenous selection of the country of residence.[9] The resulting equations have been estimated by OLS. Since our estimation strategy is a multi-step procedure, the entire process has been bootstrapped with 50 replications and bootstrapped standard errors have been used

[9] In other words, if only three choices were possible, for each individual we would compute P1, P2 and P3 the probabilities of each choice, then add to the income equation P2, P3, $P1^2$, $P2^2$, $P3^2$, P1*P2, P1*P3 and P2*P3 as independent variables. A one order polynomial could have been chosen but, as shown by Bourguignon et al. (2004), in this case Dahl's and Lee's methods yield similar results and estimates are strongly biased when the restrictions that Lee's method imposes on the data are not met. A higher order polynomial could also have been considered, but with seven possible choices the number of correction terms increases very rapidly, so a polynomial of order two appears a reasonable choice.

for hypothesis testing. Results are presented in Table 7. As the coefficients of the polynomials of the selection probabilities have no interpretation and because of space limitations, we limit the presentation to the coefficients of the variables that have a direct interpretation. The first column shows the estimated coefficients when no correction for endogenous selection is applied. The second column presents the corrected coefficients and the third column the results of a Hausman-type test for differences between the corrected and uncorrected estimated coefficients of the education, experience and language variables. The results of a series of Wald tests are also shown at the bottom of Table 7. Several test statistics were computed. First, we test whether the selection correction terms enter the earnings equation significantly. Second, we test the hypothesis that our excluded variables, that is the religion and nationality dummies, have no significant contribution to the explanation of the dependent variable, namely log-earnings. When both groups are included, identification is achieved through the non linearity of the choice probabilities (see *supra*).[10]

We find that for two countries, namely Cote d'Ivoire and Mali, we can reject the hypothesis that the coefficients of the polynomials included to correct for endogenous selection are all zero.Turning to the overidentification tests, the results allow us to conclude to the correct identification of our model: in all cases but one, Wald test statistics are found to be insignificant, indicating that the vector of variables used to instrument residential choice does not contribute to the determination of earnings, once the correction terms are included. Exception is Lome for which the null hypothesis is rejected, but at the 10% level and when only one group of identifying variables (religion dummies) is included in the

[10] All tests are based on bootstrapped standard errors.

regression.

In Benin, Senegal and Niger, and to a lesser extent in Côte d'Ivoire and Togo, correcting for endogenous selection appears to change significantly the estimated returns to education. Except for Mali, the education coefficients in the corrected equations are almost uniformly lower than the uncorrected coefficients. For example, in Benin, Senegal and Niger, the coefficient on "Foundation degree" drops respectively by 13%, 12.5% and 23% when correcting for self-selection.[11] In other words, self-selection leads to upward biases in the returns to education, implying that migrants in a given country share unobserved characteristics that make their earnings higher than average. However, the variation in returns to education between capital cities does not narrow, suggesting that country-specific amenities and other unmeasurable non-wage variables play important roles in the locational choice of individuals with different levels of education.

### 5.3 Structural model of residential choice

The last question we examine in this paper is whether earnings differentials matter in locational choice. To this end, we compute the minimum distance estimator of $\alpha$ using estimated coefficients from the reduced form multinomial logit and from the selectivity corrected earnings equations. Results are shown in Table 8, second row.[12] To assess the importance of correcting for the endogenous selection of the country of residence, we also present the estimate obtained when no correction is applied (see first row). As can be seen from the Table, correcting for endogenous selection does not dramatically change the results. In both cases,

[11] These countries' university system derives from the French system, in which, until recently, second-year students could get a diploma. We refer to it as the "foundation" degree.

[12] Standard deviations are computed using first and second stage bootstrap standard errors.

we find that individuals tend to locate in countries where their expected earnings are higher.

How robust are these results?

Since our results might be sensitive to some of the assumptions we made due to the limitations of our data, we conducted a series of robustness checks.

As a first check, we used another estimation method in the third stage of our whole procedure. Indeed, another way to recover the value of $\alpha$ is to compute unconditional earnings predictions, for each individual in each possible location, using unbiased estimates of $\beta_k$ and proceed to the estimation of the following structural conditional logit model:

$$P(i \text{ lives in } k) = P(M_i(j) = k) = \frac{\exp(\alpha \widehat{\ln y_{ik}} + z_i' \gamma_k)}{\sum_{l=1}^{P} \exp(\alpha \widehat{\ln y_{il}} + z_i' \gamma_l)} \tag{6}$$

This alternative method does not impose the linear restrictions between the coefficient values that result from the structural model (see equation 5). Results appear in Table 8, third row. While the coefficient of $\alpha$ is lower than that obtained using the minimum distance estimator, it is still significantly positive, bringing additional support to the idea that individuals tend to locate in countries where their expected earnings are higher.

As a second robustness check, in the second stage of our estimation procedure we run a Heckman selection model using data on participants and non-participants to the labour market, instead of running an OLS regression on participants only. Indeed, in the foregoing estimations, due to the difficulty of controlling both for the endogenous selection of locational choice and for labour force participation, our sample was restricted to labour market participants. This limitation is naturally a potential source of bias in our estimates.

The identifying variable in the Heckman selection model is marital status (i.e. whether the individual is married or not), which is assumed to influence labor market participation but not earnings. Results obtained in the third stage were not affected by this change, suggesting negligible biases.

Third, we checked whether self-selected *internal* migration affected the observed returns to education but found no evidence of a selection bias.

Last, since our results might depend upon the set of conversion factors used to convert current CFA francs to PPP, we re-run our model using another set of conversion factors. The latter have been computed in 1998 by ASECNA and have been actualized through 2001 using national inflation rates.[13] Once again, this modification did not change dramatically our results.

# 6 Conclusion

In this paper we use a unique set of identical labour force surveys that allow to observe at the same time migrants in seven WAEMU countries and their country of origin's labour market. We use these data first to document the patterns of migration flows in the sub-region, second to estimate the determinants of migration behaviour across these countries and finally to correct the estimated returns to education for the endogeneity of locational choice. We then use a minimum distance estimator to evaluate the impact of expected earnings differentials on the probability of selecting a particular country to reside in.

Our results show that, despite the severe political crisis that started in

[13] ASECNA is the Agence pour la Sécurité de la Navigation Aérienne en Afrique et à Madagascar. This agency computed its own PPP conversion factors, based on prices observed in the African capital cities, in order to give the same wage to its agents in terms of purchasing power.

1999, Cote d'Ivoire remains the most important immigration country in the sub-region. Our data also suggests that Mali and Burkina Faso have been and still are major labour-exporting countries, largely towards Cote d'Ivoire. Benin and Togo, by contrast, combine both emigration and immigration. Looking at migrants characteristics we find that migrants tend to be less educated than non migrants in both their origin and destination country. Thus cross-border migration within the sub-region seems to concern mainly low educated individuals. They are more likely than natives to work in the informal sector and they receive lower wages.

Our econometric results suggest that not holding account of international migration in estimating returns to education yields upward biased estimates in three countries out of seven. However, the variation in returns to education between capital cities does not narrow, suggesting that country-specific amenities and other unmeasurable non-wage variables play important roles in the locational choice of individuals with different levels of education. We also find that expected earnings differentials have a very significant effect on the choice probabilities: all else equal, people tend to live in countries in which their expected earnings are higher than elsewhere.

# References

Adebusoye, P. M. (2006), 'Geographic Labour Mobility in Sub-Saharan Africa'. IDRC Working Papers on Globalization, Growth and Poverty.

Adepoju, A. (2005), 'Migration in West Africa'. Paper prepared for the Policy Analysis and Research Programme of the Global Commision on Interna-

tional Migration.

African Union (2005), Draft Strategic Framework for a Policy on Migration in Africa, Third ordinary session of the Labour and Social Affairs Commission, 18-23 April, 2005. Pretoria, South Africa.

Amegashie, F., Brilleau, A., Coulibaly, S., Koriko, O., Ouedraogo, E., Roubaud, F. & Torelli, C. (2005), 'La conception et la mise en oeuvre des enquêtes 1-2-3 en UEMOA. Les enseignements méthodologiques.', *Stateco* **99**.

Borjas, G. (1987), 'Self selection and the earnings of migrants', *American Economic Review* **77**(4), 531–553.

Bourguignon, F., Fournier, M. & Gurgand, M. (2004), 'Selection bias corrections based on the multinomial logit model: Monte-carlo comparisons'. mimeo.

Brilleau, A., Roubaud, F. & Torelli, C. (2005), 'L'emploi, le chômage et les conditions d'activité: l'enquête 1-2-3, phase 1', *Stateco* **99**.

Dahl, G. (2002), 'Mobility and the returns to education: testing a Roy model with multiple markets', *Econometrica* **70**, 2367–2420.

Gourieroux, C. & Monfort, A. (1995), Statistics and econometric models, Cambridge University Press.

Harris, J. & Todaro, M. (1970), 'Migration, unemployment and development : a two-sector analysis', *American Economic Review* **60**, 126–142.

Heckman, J. (1979), 'Sample selection bias as a specification error', *Econometrica* **47**, 153–161.

International Organisation for Migration (2003), World Migration 2003: Challenges and Responses for People on the Move, IOM, Geneva.

Lee, L. (1983), 'Generalized econometric models with selectivity', *Econometrica* **51**, 507–512.

Roy, A. (1951), 'Some thoughts on the distribution of earnings', *Oxford Economic Papers* **3**, 135–146.

Schmertmann, C. (1994), 'Selectivity bias correction methods in polychotomous sample selection models', *Journal of Econometrics* **60**, 101–132.

Sjaastad, L. A. (1962), 'The costs and returns of human migration', *Journal of Political Economy* **70**, 80–93.

Todaro, M. (1969), 'A model of labor migration and urban unemployment in less developed countries', *American Economic Review* **59**, 138–148.

Wooldridge, J. M. (2002), Econometric analysis of cross section and panel data, MIT Press.

World Bank (2003), World Development Indicators cd-rom, World Bank, Washington DC.

**Table 1 - Composition of national samples**

| | Number of sample individuals coming from: | | | | | | | | | Total number of immigrants | Total number of natives | Total sample size |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Benin | Burkina Faso | Cote d'Ivoire | Mali | Niger | Senegal | Togo | Other | n.d. | | | |
| Benin | - | 4 | 6 | 15 | 58 | 3 | 104 | 135 | 18 | 343 | **7,030** | 7,373 |
| *nationals* (*) | - | **4** | **6** | **15** | **55** | **2** | **102** | 36 | 16 | 236 | | |
| Burkina Faso | 11 | - | 7 | 8 | 2 | 1 | 16 | 18 | 11 | 74 | **8,108** | 8,182 |
| *of which WAEMU nationals* | **6** | - | **7** | **7** | **0** | **1** | **16** | 5 | 7 | 49 | | |
| Cote d'Ivoire | 54 | 413 | - | 242 | 89 | 65 | 78 | 315 | 135 | 1391 | **5,871** | 7,262 |
| *of which WAEMU nationals* | **54** | **410** | - | **229** | **84** | **63** | **73** | 122 | 126 | 1161 | | |
| Mali | 8 | 14 | 11 | - | 8 | 14 | 0 | 64 | 7 | 126 | **7,077** | 7,203 |
| *of which WAEMU nationals* | **3** | **13** | **10** | - | **6** | **13** | **0** | 37 | 5 | 87 | | |
| Niger | 76 | 47 | 4 | 122 | - | 5 | 59 | 52 | 28 | 393 | **7,550** | 7,943 |
| *of which WAEMU nationals* | **65** | **47** | **4** | **122** | - | **5** | **48** | 27 | 25 | 343 | | |
| Senegal | 10 | 0 | 3 | 10 | 0 | - | 5 | 133 | 52 | 213 | **11,764** | 11,977 |
| *of which WAEMU nationals* | **6** | **0** | **2** | **9** | 0 | - | **2** | 74 | 34 | 127 | | |
| Togo | 87 | 9 | 9 | 11 | 51 | 3 | - | 109 | 23 | 302 | **5,873** | 6,175 |
| *of which WAEMU nationals* | **76** | **9** | **8** | **11** | **44** | **3** | - | 22 | 21 | 194 | | |
| Total | 246 | 487 | 40 | 408 | 208 | 91 | 262 | 826 | 274 | | | |
| *of which WAEMU nationals* | 210 | 483 | 37 | 393 | 189 | 87 | 241 | 323 | 234 | | | |

*Source:* 1-2-3 Surveys, 1st round, 2001-2003, National Statistical Institutes, AFRISTAT and DIAL. Authors' computations.

Note: All individuals aged 16 or more. Are considered as natives of country *i* all individuals who have always been residing in country *i*, whether they declare having the country's citizenship or not. In bold are all sample individuals who will be considered in the analysis.

(*) Within the sample of immigrants coming from one of the six WAEMU countries, some individuals are not WAEMU nationals (*Exemple*: A French national who spent 10 years in Burkina Faso before moving to Benin is recorded as an immigrant coming from Burkina Faso but is not Burkinabe).

**Table 2 - (Weighted) share of immigrants among urban residents by WAEMU country (%)**

| | Bénin | Burkina | Côte d'Ivoire | Mali | Niger | Sénégal | Togo |
|---|---|---|---|---|---|---|---|
| Natives | 96.4 | 99.3 | 84.1 | 98.4 | 95.6 | 98.5 | 95.5 |
| Immigrants | 3.6 | 0.7 | 15.9 | 1.6 | 4.4 | 1.6 | 4.5 |
| *of which:* | | | | | | | |
| coming from WAEMU | 60.6 | 70.7 | 73.5 | 43.8 | 85.7 | 13.0 | 60.7 |
| coming from other developing countries | 36.4 | 23.9 | 25.2 | 43.4 | 12.2 | 83.9 | 38.8 |
| coming from developed countries | 3.1 | 6.2 | 1.3 | 12.6 | 2.2 | 3.1 | 0.8 |

*Source:* 1-2-3 Surveys, 1st round, 2001-2003, National Statistical Institutes, AFRISTAT and DIAL. Authors' computations.

**Table 3 - Mean characteristics of natives and immigrants by country of residence**

| | Benin | | Burkina Faso | | Cote d'Ivoire | | Mali | | Niger | | Senegal | | Togo | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Natives | Immigrants | Natives | Immigrants | Natives | Immigrants | Natives | Immigrants | Natives | Immigrants | Natives | Immigrants | Natives | Immigrants |
| % of males | 48.4 | 42.4 | 51.0 | 56.8 | 48.1 | 61.8 * | 49.7 | 46.7 | 48.8 | 45.7 | 47.3 | 47.4 | 47.7 | 58.3 * |
| Age in years | 33.5 | 31.8 | 32.4 | 31.9 | 30.0 | 35.8 * | 33.4 | 33.2 | 32.6 | 35.2 * | 33.4 | 37.2 | 32.2 | 32.1 |
| **Education and experience** | | | | | | | | | | | | | | |
| Experience in years | 20.9 | 22.2 | 21.4 | 20.3 | 17.5 | 27.9 * | 22.7 | 21.5 | 21.6 | 26.9 * | 22.1 | 22.9 | 19.6 | 21.7 * |
| Years of schooling | 6.6 | 3.6 * | 5.0 | 5.6 | 6.6 | 1.9 * | 4.7 | 5.4 | 5.0 | 2.3 * | 5.2 | 8.3 * | 6.6 | 4.4 * |
| % with no diploma | 46.5 | 72.8 * | 56.2 | 54.1 | 44.7 | 84.7 * | 59.8 | 60.0 | 61.6 | 81.8 * | 61.3 | 36.8 * | 43.3 | 63.6 * |
| % with completed primary education | 25.6 | 14.7 * | 22.4 | 13.5 | 26.6 | 9.3 * | 17.5 | 13.3 | 19.0 | 11.3 * | 17.1 | 15.8 | 30.9 | 24.5 * |
| % with BEPC | 13.2 | 6.0 * | 11.4 | 18.9 | 10.6 | 2.4 * | 8.0 | 4.4 | 7.4 | 2.4 * | 11.1 | 15.8 | 14.8 | 4.6 * |
| % with baccalaureat | 4.0 | 3.8 | 2.7 | 0.0 | 4.9 | 0.7 * | 2.2 | 6.7 * | 2.7 | 0.0 * | 3.9 | 5.3 | 3.2 | 1.3 |
| Can read&write in French | 70.7 | 36.4 * | 57.3 | 62.2 | 73.2 | 26.8 * | 47.6 | 46.7 | 54.8 | 29.6 * | 58.7 | 73.7 | 72.5 | 52.3 * |
| Can read&write in a foreign language | 24.1 | 26.1 | 13.1 | 27.0 * | 24.9 | 10.8 * | 12.3 | 33.3 * | 21.4 | 18.2 | 19.1 | 42.1 * | 26.7 | 23.2 |
| **Religion** | | | | | | | | | | | | | | |
| % of muslim | 9.7 | 46.2 * | 56.2 | 40.5 * | 31.0 | 73.6 * | 97.3 | 80.0 * | 98.3 | 77.0 * | 93.2 | 57.9 * | 9.4 | 49.0 * |
| % of catholic | 67.8 | 32.1 * | 35.9 | 16.2 * | 35.8 | 17.4 * | 1.7 | 17.8 * | 1.1 | 18.9 * | 6.6 | 42.1 * | 47.9 | 22.5 * |
| % of protestant | 5.1 | 3.8 | 6.5 | 27.0 * | 10.8 | 3.5 * | 0.5 | 2.2 | 0.4 | 3.4 * | 0.1 | 0.0 | 10.3 | 1.3 * |
| **Number of observations** | 7,030 | 184 | 8,108 | 37 | 5,871 | 913 | 7,077 | 45 | 7,550 | 291 | 11,764 | 19 | 5,873 | 151 |

*Source:* 1-2-3 Surveys, 1st round, 2001-2003, National Statistical Institutes, AFRISTAT and DIAL. Authors' computations.

A "*" means that the difference is statistically significant

**Table 4 - Mean characteristics of natives and emigrants by country of residence**

| | Benin | | Burkina Faso | | Cote d'Ivoire | | Mali | | Niger | | Senegal | | Togo | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Natives | Emigrants | Natives | Emigrants | Natives | Emigrants | Natives | Emigrants | Natives | Emigrants | Natives | Emigrants | Natives | Emigrants |
| % of males | 48.4 | 46.7 | 51.0 | 58.8 * | 48.1 | 51.4 | 49.7 | 58.0 * | 48.8 | 68.3 * | 47.3 | 70.1 * | 47.7 | 39.4 * |
| Age in years | 33.5 | 34.6 * | 32.4 | 36.1 * | 30.0 | 29.4 | 33.4 | 36.3 * | 32.6 | 32.0 | 33.4 | 38.2 * | 32.2 | 31.4 |
| **Education and experience** | | | | | | | | | | | | | | |
| Experience in years | 20.9 | 23.4 * | 21.4 | 28.6 * | 17.5 | 17.1 | 22.7 | 29.1 * | 21.6 | 23.7 * | 22.1 | 27.8 * | 19.6 | 21.3 * |
| Years of schooling | 6.6 | 5.2 * | 5.0 | 1.6 * | 6.6 | 6.3 | 4.7 | 1.1 * | 5.0 | 2.3 * | 5.2 | 4.1 * | 6.6 | 4.1 * |
| % with no diploma | 46.5 | 56.2 * | 56.2 | 87.8 * | 44.7 | 45.9 | 59.8 | 91.1 * | 61.6 | 81.5 * | 61.3 | 66.7 | 43.3 | 68.9 * |
| % with completed primary education | 25.6 | 23.3 | 22.4 | 8.1 * | 26.6 | 18.9 | 17.5 | 6.1 * | 19.0 | 11.6 * | 17.1 | 13.8 | 30.9 | 17.8 * |
| % with BEPC | 13.2 | 7.6 * | 11.4 | 1.7 * | 10.6 | 8.1 | 8.0 | 1.0 * | 7.4 | 2.6 * | 11.1 | 6.9 | 14.8 | 7.1 * |
| % with baccalaureat | 4.0 | 1.0 * | 2.7 | 0.4 * | 4.9 | 8.1 | 2.2 | 0.3 * | 2.7 | 1.6 | 3.9 | 6.9 | 3.2 | 0.8 * |
| Can read&write in French | 70.7 | 56.2 * | 57.3 | 25.5 * | 73.2 | 62.2 | 47.6 | 15.8 * | 54.8 | 27.5 * | 58.7 | 50.6 | 72.5 | 46.9 * |
| Can read&write in a foreign language | 24.1 | 19.5 | 13.1 | 8.1 * | 24.9 | 37.8 * | 12.3 | 13.2 | 21.4 | 32.1 * | 19.1 | 21.8 | 26.7 | 17.4 * |
| **Religion** | | | | | | | | | | | | | | |
| % of muslim | 9.7 | 26.2 * | 56.2 | 70.6 * | 31.0 | 51.4 * | 97.3 | 99.0 * | 98.3 | 96.3 | 93.2 | 86.2 * | 9.4 | 23.2 * |
| % of catholic | 67.8 | 38.6 * | 35.9 | 25.1 * | 35.8 | 16.2 * | 1.7 | 0.3 * | 1.1 | 1.6 | 6.6 | 10.3 | 47.9 | 44.8 |
| % of protestant | 5.1 | 7.6 | 6.5 | 2.3 * | 10.8 | 2.7 | 0.5 | 0.8 | 0.4 | 0.5 | 0.1 | 1.1 * | 10.3 | 12.0 |
| **Number of observations** | 7,030 | 210 | 8,108 | 483 | 5,871 | 37 | 7,077 | 393 | 7,550 | 189 | 11,764 | 87 | 5,873 | 241 |

*Source:* 1-2-3 Surveys, 1st round, 2001-2003, National Statistical Institutes, AFRISTAT and DIAL. Authors' computations.

A "*" means that the difference is statistically significant

**Table 5 - Employment situation of natives and immigrants, by country of residence**

| | Benin | | Burkina Faso | | Cote d'Ivoire | | Mali | | Niger | | Senegal | | Togo | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Natives | Immigrants | Natives | Immigrants | Natives | Immigrants | Natives | Immigrants | Natives | Immigrants | Natives | Immigrants | Natives | Immigrants |
| **Employment situation** | | | | | | | | | | | | | | |
| % of employed | 67.5 | 70.1 | 57.1 | 56.8 | 60.7 | 77.7 | 57.6 | 55.6 | 48.4 | 66.3 | 50.1 | 57.9 | 70.5 | 77.5 |
| % of unemployed | 4.1 | 2.2 | 11.0 | 18.9 | 11.5 | 4.7 | 4.2 | 2.2 | 8.0 | 3.8 | 7.5 | 0.0 | 6.8 | 4.0 |
| % of inactive | 28.4 | 27.7 | 31.9 | 24.3 | 27.8 | 17.6 | 38.1 | 42.2 | 43.6 | 29.9 | 42.4 | 42.1 | 22.7 | 18.5 |
| Number of observations | 7,030 | 184 | 8,108 | 37 | 5,871 | 913 | 7,077 | 45 | 7,550 | 291 | 11,764 | 19 | 5,873 | 151 |
| **Sector of activity & wage of the employed** | | | | | | | | | | | | | | |
| % in the public sector | 8.9 | 0.0 | 14.0 | 9.5 | 8.4 | 1.0 | 11.6 | 4.0 | 18.0 | 1.0 | 9.1 | 0.0 | 8.2 | 1.7 |
| % in the formal private sector | 11.7 | 10.9 | 9.1 | 19.1 | 21.5 | 12.6 | 11.8 | 8.0 | 13.5 | 10.4 | 17.9 | 36.4 | 8.3 | 12.0 |
| % in the informal private sector | 79.4 | 89.1 | 76.9 | 71.4 | 70.1 | 86.5 | 76.6 | 88.0 | 68.5 | 88.6 | 73.0 | 63.6 | 83.6 | 86.3 |
| Hourly wage in PPP CFA Francs | 200 | 144 | 384 | 369 | 356 | 228 | 340 | 434 | 487 | 344 | 421 | 1,090 | 309 | 428 |
| Number of observations | 4,745 | 129 | 4,630 | 21 | 3,564 | 709 | 4,076 | 25 | 3,654 | 193 | 5,894 | 11 | 4,140 | 117 |

*Source:* 1-2-3 Surveys, 1st round, 2001-2003, National Statistical Institutes, AFRISTAT and DIAL. Authors' computations.

**Table 6 - Results of the reduced form multinomial logit**

| | Cotonou (Benin) | Ouagadougou (Burkina Faso) | Abidjan (Côte d'Ivoire) | Bamako (Mali) | Niamey (Niger) | Lome (Togo) |
|---|---|---|---|---|---|---|
| Sex (1: Male) | 0.96*** | 0.79** | 1.36*** | 0.24 | 0.37 | 1.13*** |
| | (0.33) | (0.31) | (0.29) | (0.32) | (0.32) | (0.33) |
| CEP (Primary school completed) | -0.63 | 0.45 | -0.20 | -0.01 | -0.34 | -0.05 |
| | (0.50) | (0.49) | (0.45) | (0.50) | (0.50) | (0.49) |
| BEPC (GCSE) | -0.69 | 1.17* | 0.04 | 0.33 | -0.47 | -0.12 |
| | (0.64) | (0.67) | (0.57) | (0.70) | (0.66) | (0.64) |
| CAP | -0.73 | 1.95* | 0.83 | 2.01* | 0.20 | -0.76 |
| | (1.13) | (1.16) | (1.03) | (1.15) | (1.16) | (1.13) |
| BEP | -2.96 | 1.27 | -0.88 | 0.66 | -0.68 | -0.05 |
| | (1.90) | (1.96) | (1.81) | (1.80) | (1.85) | (1.85) |
| Baccalaureate | 1.50* | 2.54*** | 1.32** | 1.57* | 1.57** | 1.87** |
| | (0.79) | (0.88) | (0.60) | (0.86) | (0.78) | (0.77) |
| Foundation degree | -2.18 | 1.40 | -0.08 | 0.57 | -0.53 | -1.00 |
| | (1.70) | (1.81) | (1.66) | (1.75) | (1.75) | (1.75) |
| Bachelor's degree | -1.99* | 0.32 | -0.99 | 0.04 | -0.57 | -1.65 |
| | (1.09) | (1.14) | (1.03) | (1.10) | (1.09) | (1.10) |
| Postgraduate degree | -5.01*** | -3.65*** | -4.39*** | -3.51*** | -3.57*** | -4.92*** |
| | (0.97) | (1.14) | (1.02) | (1.05) | (1.02) | (1.03) |
| Marital status (1: Married) | -0.45 | -0.70** | -0.89*** | -0.14 | -0.50 | -0.53 |
| | (0.34) | (0.33) | (0.30) | (0.33) | (0.33) | (0.34) |
| Speaks French (1=Yes) | -0.14 | 0.00 | -0.20 | -0.02 | 0.29 | 0.25 |
| | (0.39) | (0.36) | (0.33) | (0.38) | (0.38) | (0.38) |
| Speaks another Foreign Language (1=Yes) | 1.01** | -0.04 | 0.10 | -0.21 | 0.11 | 0.92** |
| | (0.41) | (0.40) | (0.36) | (0.39) | (0.40) | (0.40) |
| Experience (in years) | 0.06 | 0.05 | 0.16*** | -0.00 | 0.10** | 0.07 |
| | (0.04) | (0.04) | (0.04) | (0.04) | (0.04) | (0.04) |
| Experience Squared | -0.00 | -0.00 | -0.00*** | 0.00 | -0.00 | -0.00 |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Public sector | -0.07 | 0.58 | -1.53*** | 0.30 | -0.09 | -0.23 |
| | (0.60) | (0.64) | (0.57) | (0.61) | (0.60) | (0.60) |
| Private sector | 0.12 | -0.86** | -0.18 | 0.26 | 0.16 | -0.27 |
| | (0.42) | (0.39) | (0.35) | (0.40) | (0.40) | (0.42) |
| Father in the agricultural sector | -0.27 | -0.06 | 0.44 | -0.51 | -0.17 | -0.10 |
| | (0.36) | (0.34) | (0.32) | (0.34) | (0.35) | (0.35) |
| Father in the industrial sector | -0.81 | -0.59 | 0.04 | -0.38 | -0.29 | -0.63 |
| | (0.57) | (0.59) | (0.49) | (0.57) | (0.58) | (0.57) |
| Father in the commercial sector | 0.61 | 1.16*** | 1.20*** | 0.89** | 0.52 | 1.01** |
| | (0.40) | (0.38) | (0.32) | (0.37) | (0.38) | (0.40) |
| Father was a top executive | 0.79 | 2.43*** | 1.55** | 1.65** | 1.64** | 1.10 |
| | (0.74) | (0.81) | (0.63) | (0.75) | (0.75) | (0.74) |
| Father was a middle executive | 1.17** | 0.97 | 0.75 | 1.18* | 1.24** | 0.56 |
| | (0.60) | (0.63) | (0.54) | (0.61) | (0.60) | (0.59) |
| Father never went to school | 1.20*** | 1.85*** | 1.17*** | 0.59* | 1.53*** | 0.74** |
| | (0.33) | (0.32) | (0.29) | (0.32) | (0.33) | (0.33) |
| Muslim | -4.44*** | -3.98*** | -4.17*** | -4.15*** | -1.57 | -5.29*** |
| | (1.31) | (1.35) | (1.29) | (1.40) | (1.37) | (1.31) |
| Catholic | -2.68** | -3.03** | -3.41** | -3.48** | -1.69 | -3.86*** |
| | (1.34) | (1.38) | (1.33) | (1.45) | (1.40) | (1.34) |
| Protestant | -1.31 | 0.33 | -0.66 | -1.71 | 0.38 | -2.13 |
| | (1.76) | (1.79) | (1.73) | (1.88) | (1.82) | (1.75) |
| Nationality dummies | *Included but not shown* | | | | | |
| Intercept | -7.03*** | -8.12*** | -4.65*** | -3.96** | -9.08*** | -5.49*** |
| | (1.80) | (1.85) | (1.46) | (1.61) | (1.66) | (1.59) |
| Observations | 31,426 | 31,426 | 31,426 | 31,426 | 31,426 | 31,426 |

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

**Table 7 - Earnings regression - uncorrected (1st column) and corrected (2nd column) estimates**

| | Cotonou (Benin) | | | Ouagadougou (Burkina Faso) | | | Abidjan (Côte d'Ivoire) | | | Bamako (Mali) | | | Niamey (Niger) | | | Dakar (Senegal) | | | Lome (Togo) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sex (1: Male) | 0.46*** | 0.45*** | | 0.42*** | 0.26*** | | 0.42*** | 0.41*** | | 0.35*** | 0.22*** | | 0.23*** | 0.16* | | -0.10** | -0.20** | | 0.34*** | 0.29*** | |
| | (0.04) | (0.07) | | (0.05) | (0.09) | | (0.05) | (0.05) | | (0.04) | (0.07) | | (0.05) | (0.09) | | (0.05) | (0.08) | | (0.05) | (0.07) | |
| CEP (Primary school completed) | 0.50*** | 0.50*** | 0.04 | 0.40*** | 0.33*** | 1.11 | 0.54*** | 0.54*** | 0.27 | 0.22*** | 0.23** | 0.06 | 0.52*** | 0.45*** | 1.48 | 0.35*** | 0.32*** | 2.61 | 0.47*** | 0.47*** | 0.05 |
| | (0.06) | (0.07) | | (0.08) | (0.14) | | (0.07) | (0.07) | | (0.07) | (0.11) | | (0.09) | (0.13) | | (0.07) | (0.10) | | (0.07) | (0.09) | |
| BEPC (GCSE) | 0.96*** | 0.92*** | 1.24 | 1.27*** | 1.23*** | 0.17 | 1.16*** | 1.16*** | 0.22 | 0.47*** | 0.54*** | 0.91 | 1.06*** | 0.92*** | 4,57** | 0.55*** | 0.45*** | 9,18*** | 1.05*** | 1.02*** | 0.54 |
| | (0.09) | (0.10) | | (0.11) | (0.17) | | (0.10) | (0.10) | | (0.11) | (0.14) | | (0.13) | (0.16) | | (0.09) | (0.13) | | (0.10) | (0.13) | |
| CAP | 1.15*** | 0.83*** | 8,92*** | 1.17*** | 1.08*** | 0.55 | 1.21*** | 1.21*** | 0.02 | 0.45*** | 0.57*** | 0.74 | 1.49*** | 1.16*** | 3,20* | 0.71*** | 0.22 | 8,83*** | 1.06*** | 0.89*** | 2,74* |
| | (0.15) | (0.25) | | (0.18) | (0.20) | | (0.20) | (0.15) | | (0.12) | (0.22) | | (0.23) | (0.26) | | (0.24) | (0.39) | | (0.23) | (0.25) | |
| BEP | 0.97** | 0.83 | 0.15 | 1.71*** | 1.58*** | 1.05 | 1.15*** | 1.07*** | 2,80* | 0.95*** | 1.04*** | 0.65 | 1.31*** | 1.01*** | 1.78 | 0.83*** | 0.66** | 8,05*** | 1.23*** | 1.27*** | 0.13 |
| | (0.47) | (0.83) | | (0.25) | (0.32) | | (0.19) | (0.15) | | (0.11) | (0.18) | | (0.20) | (0.29) | | (0.26) | (0.24) | | (0.23) | (0.18) | |
| Baccalaureate | 1.33*** | 1.36*** | 0.66 | 1.81*** | 1.79*** | 0.06 | 1.69*** | 1.70*** | 0.04 | 0.75*** | 0.76*** | 0.01 | 1.86*** | 1.90*** | 0.68 | 0.92*** | 1.01*** | 0.39 | 1.57*** | 1.56*** | 0.01 |
| | (0.15) | (0.12) | | (0.19) | (0.17) | | (0.15) | (0.14) | | (0.20) | (0.27) | | (0.19) | (0.23) | | (0.14) | (0.29) | | (0.17) | (0.19) | |
| Foundation degree | 2.03*** | 1.77*** | 4,68** | 2.06*** | 1.97*** | 0.55 | 2.04*** | 2.01*** | 0.64 | 0.96*** | 1.10*** | 1.57 | 1.87*** | 1.44*** | 5,32** | 1.12*** | 0.98*** | 7,98*** | 2.65*** | 2.58*** | 0.79 |
| | (0.20) | (0.32) | | (0.22) | (0.32) | | (0.15) | (0.15) | | (0.16) | (0.19) | | (0.25) | (0.33) | | (0.25) | (0.19) | | (0.28) | (0.20) | |
| Bachelor's degree | 1.91*** | 1.76*** | 6,07** | 2.36*** | 2.30*** | 0.33 | 2.27*** | 2.23*** | 1.15 | 1.36*** | 1.50*** | 2.09 | 2.25*** | 2.02*** | 3,41* | 1.36*** | 1.30*** | 1.92 | 2.45*** | 2.47*** | 0.17 |
| | (0.13) | (0.16) | | (0.16) | (0.19) | | (0.13) | (0.12) | | (0.12) | (0.14) | | (0.14) | (0.20) | | (0.14) | (0.16) | | (0.16) | (0.14) | |
| Postgraduate degree | 1.68*** | 1.23*** | 7,42*** | 1.56*** | 1.49*** | 0.68 | 1.80*** | 1.83*** | 0.29 | 1.11*** | 1.11 | 0 | 1.99*** | 1.82*** | 2.32 | 1.36*** | 1.33*** | 0.29 | 2.05*** | 1.92*** | 0.56 |
| | (0.18) | (0.28) | | (0.23) | (0.29) | | (0.21) | (0.19) | | (0.23) | (0.74) | | (0.17) | (0.20) | | (0.18) | (0.17) | | (0.26) | (0.33) | |
| Marital status (1: Married) | 0.66*** | 0.69*** | | 0.38*** | 0.46*** | | 0.28*** | 0.28*** | | 0.46*** | 0.49*** | | 0.50*** | 0.55*** | | 0.37*** | 0.41*** | | 0.52*** | 0.54*** | |
| | (0.05) | (0.06) | | (0.06) | (0.08) | | (0.05) | (0.04) | | (0.05) | (0.07) | | (0.06) | (0.08) | | (0.05) | (0.07) | | (0.06) | (0.07) | |
| Speaks French (1: Yes) | 0.14** | 0.21*** | 11,27*** | 0.42*** | 0.46*** | 0.89 | 0.07 | 0.09 | 2.58 | 0.23*** | 0.22** | 0.14 | 0.26*** | 0.34*** | 2 | 0.30*** | 0.35*** | 8,10*** | 0.06 | 0.11 | 3,65* |
| | (0.06) | (0.08) | | (0.07) | (0.10) | | (0.06) | (0.07) | | (0.06) | (0.09) | | (0.07) | (0.11) | | (0.06) | (0.07) | | (0.07) | (0.08) | |
| Speaks a foreign language (1: Yes | 0.35*** | 0.44*** | 6,55** | 0.33*** | 0.34*** | 0.03 | 0.18** | 0.19** | 0.01 | 0.14** | -0.01 | 4,82** | 0.08 | 0.26** | 4,14** | 0.35*** | 0.40*** | 1.97 | 0.03 | 0.02 | 0.15 |
| | (0.07) | (0.08) | | (0.08) | (0.10) | | (0.07) | (0.09) | | (0.06) | (0.10) | | (0.07) | (0.11) | | (0.07) | (0.09) | | (0.07) | (0.10) | |
| Experience (in years) | 0.13*** | 0.13*** | 1.86 | 0.13*** | 0.12*** | 2.13 | 0.10*** | 0.10*** | 0.36 | 0.07*** | 0.06*** | 1.98 | 0.13*** | 0.12*** | 1.89 | 0.12*** | 0.11*** | 1.59 | 0.13*** | 0.13*** | 0.16 |
| | (0.01) | (0.01) | | (0.01) | (0.01) | | (0.01) | (0.01) | | (0.01) | (0.01) | | (0.01) | (0.01) | | (0.01) | (0.01) | | (0.01) | (0.01) | |
| Experience squared | -0.00*** | -0.00*** | 3,17* | -0.00*** | -0.001*** | 2.01 | -0.00*** | -0.001*** | 0.1 | -0.00*** | -0.001*** | 1.44 | -0.00*** | -0.001*** | 2.69 | -0.00*** | -0.001*** | 4,92* | -0.00*** | -0.001*** | 0.47 |
| | (0.00) | (0.00) | | (0.00) | (0.00) | | (0.00) | (0.00) | | (0.00) | (0.00) | | (0.00) | (0.00) | | (0.00) | (0.00) | | (0.00) | (0.00) | |
| Public sector | 0.30*** | 0.41*** | | 0.69*** | 0.83*** | | 0.72*** | 0.63*** | | 0.34*** | 0.45*** | | 0.49*** | 0.63*** | | 0.80*** | 0.88*** | | 0.68*** | 0.68*** | |
| | (0.08) | (0.11) | | (0.08) | (0.16) | | (0.10) | (0.09) | | (0.07) | (0.10) | | (0.08) | (0.12) | | (0.09) | (0.10) | | (0.10) | (0.09) | |
| Private sector | 0.33*** | 0.32*** | | 0.44*** | 0.54*** | | 0.66*** | 0.66*** | | 0.18*** | 0.26** | | 0.31*** | 0.37*** | | 0.80*** | 0.83*** | | 0.33*** | 0.34*** | |
| | (0.07) | (0.08) | | (0.09) | (0.12) | | (0.06) | (0.06) | | (0.06) | (0.12) | | (0.08) | (0.11) | | (0.06) | (0.08) | | (0.09) | (0.08) | |
| Father in the agricultural sector | -0.01 | -0.07 | | -0.18*** | -0.32*** | | -0.09* | -0.10* | | -0.12** | -0.22*** | | -0.03 | -0.17* | | 0.01 | -0.06 | | 0.05 | 0.03 | |
| | (0.05) | (0.08) | | (0.05) | (0.08) | | (0.05) | (0.06) | | (0.05) | (0.08) | | (0.06) | (0.09) | | (0.06) | (0.07) | | (0.06) | (0.06) | |
| Father in the industrial sector | 0.13 | 0.03 | | -0.30* | -0.36 | | -0.19** | -0.20 | | -0.05 | -0.07 | | -0.15 | -0.24 | | -0.11 | -0.13 | | -0.08 | -0.08 | |
| | (0.09) | (0.12) | | (0.15) | (0.27) | | (0.09) | (0.13) | | (0.09) | (0.13) | | (0.14) | (0.23) | | (0.07) | (0.09) | | (0.11) | (0.11) | |
| Father in the commercial sector | 0.08 | 0.05 | | 0.03 | -0.08 | | -0.03 | -0.05 | | 0.07 | 0.06 | | -0.10 | -0.23 | | 0.04 | -0.11 | | 0.10 | 0.04 | |
| | (0.07) | (0.10) | | (0.08) | (0.10) | | (0.07) | (0.09) | | (0.05) | (0.09) | | (0.08) | (0.14) | | (0.07) | (0.10) | | (0.09) | (0.9) | |
| Father was a top executive | 0.26** | 0.19 | | 0.18 | 0.22 | | 0.38*** | 0.39** | | 0.41*** | 0.42*** | | -0.14 | -0.31 | | 0.25** | 0.09 | | 0.13 | 0.11 | |
| | (0.12) | (0.20) | | (0.16) | (0.21) | | (0.14) | (0.18) | | (0.10) | (0.16) | | (0.16) | (0.24) | | (0.13) | (0.21) | | (0.16) | (0.20) | |
| Father was a middle executive | 0.22*** | 0.22** | | 0.09 | 0.16 | | -0.06 | -0.09 | | 0.12* | 0.14 | | -0.03 | -0.00 | | 0.11 | 0.10 | | -0.05 | -0.04 | |
| | (0.07) | (0.09) | | (0.12) | (0.15) | | (0.09) | (0.10) | | (0.07) | (0.12) | | (0.11) | (0.14) | | (0.09) | (0.15) | | (0.09) | (0.09) | |
| Father never went to school | -0.01 | -0.00 | | -0.05 | 0.05 | | 0.06 | 0.09* | | -0.03 | -0.05 | | 0.05 | 0.08 | | 0.01 | 0.04 | | -0.19*** | -0.18** | |
| | (0.05) | (0.06) | | (0.07) | (0.15) | | (0.05) | (0.05) | | (0.04) | (0.11) | | (0.07) | (0.12) | | (0.05) | (0.10) | | (0.05) | (0.07) | |
| Intercept | 1.04*** | | | 1.58*** | | | 2.16*** | | | 3.01*** | | | 1.91*** | | | 2.34*** | | | 1.88*** | | |
| | (0.09) | | | (0.11) | | | (0.10) | | | (0.08) | | | (0.12) | | | (0.09) | | | (0.10) | | |
| Observations | 4,723 | | | 4,453 | | | 4,188 | | | 4,022 | | | 3,686 | | | 5,383 | | | 4,193 | | |
| R-squared | 0.41 | | | 0.39 | | | 0.40 | | | 0.31 | | | 0.38 | | | 0.32 | | | 0.33 | | |

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

| | | | | | | |
|---|---|---|---|---|---|---|
| Wald test for selection correction terms | 41.0 | 23.6 | 49.1** | 51.88** | 36.70 | 22.34 | 26.14 |
| Over-identification Wald Tests | | | | | | | |
| One group added | | | | | | | |
| - Religion dummies | 5.89 | 2.90 | 2.65 | 1.27 | 2.72 | 0.20 | 9.00* |
| - Nationality dummies | 2.00 | 0.04 | 2.91 | 0.79 | 1.54 | 1.04 | 4.57 |
| Both groups added | 6.36 | 2.92 | 4.58 | 1.53 | 4.18 | 1.00 | 10.64 |

**Table 8 - Results of third step minimum distance estimation**

| | Estimated value of $\alpha$ |
|---|---|
| Uncorrected model | 2.03***<br>(0.21) |
| Corrected model | 2.50***<br>(0.13) |
| Conditional logit estimation | 0.81***<br>(0.15) |

Standard deviations are reported between parentheses.
***: significant at the 1% level.