Revue-IRS

Revue Internationale de la Recherche Scientifique
(Revue-IRS)
ISSN: 2958-8413
Vol. 2, No. 4, Septembre 2024


# Analysis of the impact of education on poverty reduction in Kinshasa and Lubumbashi

**Mwamba Kande Franklin[a,b], Oshasha Oshasha Fiston[a,c], Biaba Kuya Jirince[a,d], Muhindo Rugishi Guy[a], Frey Sylvestre[a,e,f], Iwazie Pelepoko Jonathan[a,g]**

a. University of Kinshasa (UNIKIN), P.O. Box. 15.373 , Democratic Republic of Congo.
b. Institute of Health Sciences Research, P.O. Box 10.183, Democratic Republic of Congo
c. General Commissariat for Atomic Energy/Regional Center for Nuclear Studies of Kinshasa, P.O. Box 868, University of Kinshasa Campus/Democratic Republic of Congo
d. Institut Francophone International (IFI)/ Université Nationale du Vietnam, Hanoi, adresse 144 Xuan Thuy, Cau Giay, Hanoi, Vietnam
e. Forest Management Technique, Regional Postgraduate School of Integrated Management and Management of Forests and Tropical Territories (ÉRAIFT), P.O. Box 15.373, University of Kinshasa, Democratic Republic of Congo.
f. Information system and AI, Green Computer Science, Faculty of Science and Technology.
g. National Institute of Statistics

**Abstract:** Almost every country in the world suffers from poverty, and the DRC is not left out either (Sachs 2005). This paper questions the role of education in improving the condition of households in Kinshasa and Lubumbashi, hence contributing knowledge to reduction policies on this problem.

We categorize poor and non-poor households using results from the 1-2-3 2005 survey by INS (INS 2005). In Kinshasa, the poverty incidence is 42.53%, while the depth is 13.04% and the severity is 5.82%. For Lubumbashi, these are 66.86%, 34.55%, and 20.69%, respectively. Thus, financial difficulties cause 28.88% of the children dropping out of school in Kinshasa compared to their percentage share of 54.17% in Lubumbashi. In this city, poverty is more severe.

This descriptive analysis shows that poorly educated persons in Lubumbashi account for 89.55% of the population, while in Kinshasa, the percentage is 55.56%. While in Kinshasa, the poverty rate is 16.34%, in the case of Lubumbashi, it is 26.53% for people who got a university education.

Logistic regression results portray getting a degree as the ideal path to avoiding poverty.

**Keywords**: Logistic regression; education; poverty; Kinshasa; Lubumbashi, 1-2-3 survey.

**Digital Object Identifier (DOI)**: https://doi.org/10.5281/zenodo.13764921

## 1 Introduction

The following Nelson Mandela quote applies very appropriately while elaborating on the essential role of education in the socioeconomic development of a nation: “Education is the most powerful weapon you can use to change the world" (Mandela 2003). Most of the studies at a macroeconomic level confirm that education is one of the major determinant factors of GDP growth for any country, as it improves human capital (Becker 1993; Hanushek and Woessmann 2008).

Etymologically, education means "to lead out of" or "to develop." It is conceived as acquiring and developing a set of intellectual, moral, physical, and scientific knowledge and capabilities regarded as essential to attain a level of culture adjusted to a given historical context. In the Democratic Republic of Congo, education aims at training people who can transform their environment, work with love, and guide learners in choosing their specialization.

Socio-economic conditions remain precarious in the DRC. The Congolese population is among the poorest in the world, living in a prosperous country. According to the IMF Publication of October 2015, 80% of Congolese live below USD 1.25 per day (IMF 2015). For Kinshasa, this translates to 42.53% of households living below the monetary poverty line. This finding exposes the educational system's failures, inability to attain its set objectives, and worsened by the high cost of it, reducing its accessibility and making many a head of households ignore the education of their children. Only 41% of Congolese have gone to secondary school, resulting in a low enrollment and graduation rate. Seventy percent of the children do not attend school due to weak finances from their parents (UNESCO 2023).

Given this sad situation, our study sets out to assess the role of education in reducing poverty levels in both Kinshasa and Lubumbashi. In this project, we will estimate how educational qualifications influence the probability of a household being poor using logistic regression. We will need monetary poverty measures concerning the level of education and the status of poverty in these two towns. Our ultimate work will result in building education policy and enhancing the poverty reduction strategies in DRC.

## 2 Methodology

This scientific study primarily aims to assess the role of education in reducing poverty in both Kinshasa and Lubumbashi. Concerning this, we intend to establish a link between the general level of education of households and their status concerning poverty using logistic regression since the variable that predicates poverty is qualitative and binary. This chapter briefly describes the basic concepts of logistic regression, which shall be helpful in analyzing subsequent importance. We will further introduce the confusion matrix, both essential when assessing the quality of the logistic model. Finally, we will make a short presentation of the data at our disposal.

### 2.1 Data Source

The 1-2-3 survey, developed by a DIAL researcher, all aims at the analysis of the informal sector. It was applied for the first time in Yaoundé, Cameroon, from 1994-1995. Given its efficiency, since 1997, Afristat has recommended its generalization at the world level. Many countries on three continents use this survey (Afristat 1997):

a. **Africa:**
    - Cameroon
    - Democratic Republic of Congo
    - UEMOA countries
    - Morocco
    - Gabon
    - Burundi

b. **Latin America:**
    - Guatemala
    - Peru
    - Haiti

c. **Asia:**
    - Philippines
    - Mongolia
    - China
    - Bangladesh

The survey is called 1-2-3 because it is divided into three phases:

d. Phase 1: Measurement of employment and socio-economic and demographic characteristics.
e. Phase 2: Study of informal production units.
f. Phase 3: Analysis of household consumption expenditures.

We also refer to it, because of this segmentation, as the 1-2-3 surveys.

In the DRC, these surveys were conducted twice: in 2005 and 2012. The second one in 2012 was criticized by congolese scientists, due to controversial results with a supposed decrease of the poverty rate in Kinshasa from 42.5% to 26.1% between the 2005 and 2012 surveys, respectively, which is a huge drop, not at all reflected in reality. Due to technical considerations that the problems observed in the 2012 survey are not resolved, we decided to use the data from the 2005 survey for this study.

### 2.2 Evaluation of Household Expenditures on Education

Most the households are convinced that education forms the foundation of their children's future. This is why they devote a significant portion of their income to the education of their children. Therefore, we, as part of this study, will first assess household expenditures for.

### 2.3 Measurement of Poverty

Our study seeks to review the relationship between education and poverty. In this respect, to achieve this, we need an understanding of the measurement of monetary poverty.

#### 2.3.1 Poverty Indicators:

Income or consumption expenditure.

#### 2.3.2 Poverty Threshold:

- **Absolute:** Calculated based on essential needs, such as the nutritional approach.
- **Relative:** expressed in terms of percentage of the median income, usually 50 percent, 60 percent, or 70 percent. This threshold makes poverty entirely subjective.

#### 2.3.3 Poverty Indices: These respect ethical and moral axioms, such as those proposed by Sen:

- **Monotonicity Axiom:** A decrease in income below the poverty threshold increases poverty.
- **Transfer Axiom:** A transfer of income from the poor to the rich increases poverty.
- **Focus Axiom:** An increase in income above the poverty threshold does not affect overall poverty.

In this study, we will use the H, I, and FGT indices. The following elements are defined:

- **N:** The total population size
- **p:** The number of poor individuals
- **Z:** The value of the poverty threshold
- $r_i$ : the income of the $i^{th}$ individual with$1 \leq i \leq p \ et \ r_{j-1} \leq r_j \leq r_{j+1}, j = 2 \dots p-1$

##### 2.3.3.1 H Index

The H index, or "poverty rate," measures the proportion of poor individuals in the population:

$$\boldsymbol{H = \frac{p}{N}}$$

It is the most accessible and most commonly used index. However, it does not distinguish between differences in well-being amongst people with low incomes, disregarding Sen's axioms. A policy set by this index tries to finance those near the poverty line but leaves behind those at the very bottom of the poverty line.

#### 2.3.3.2 I Index

The I Index, or poverty depth index, measures the average gap between the income of the poor and the poverty threshold:

$$\boldsymbol{I} = \frac{\frac{\mathbf{1}}{\boldsymbol{N}}\sum_{i=1}^{p} \boldsymbol{d_i}}{\boldsymbol{Z}} = \frac{\mathbf{1}}{\boldsymbol{NZ}}\sum_{i=1}^{p} \boldsymbol{d_i}$$

where $d_i = Z - r_i$.. This index respects the monotonicity axiom but not the transfer axiom, as it does not consider the distribution of income among the poor.

#### 2.3.3.3 FGT Index

The FGT index (Foster, Greer, and Thorbecker) is decomposable and measures poverty by taking subgroups into account:

$$\boldsymbol{FGT_\alpha} = \frac{\mathbf{1}}{\boldsymbol{N}}\sum_{i=1}^{p} \left(\frac{\boldsymbol{d_i}}{\boldsymbol{z}}\right)^{\alpha}$$

where α is a parameter that measures the importance given to the poorest individuals in the population. The higher the value of α, the greater the emphasis on the poor. It is, therefore, an indicator of poverty aversion according to the authors.

It is important to note that this index satisfies the transfer axiom for values of α > 1 and the sensitivity axiom for values of α > 2. Finally, it quickly follows that the I and H indices are particular cases of the FGT index.:

- Si $\alpha = 0$ alors $FGT_0 = H$
- Si $\alpha = 1$ alors $FGT_1 = I$.

Furthermore:

- If α=2, then $\boldsymbol{FGT_2} = \frac{1}{N}\sum_{i=1}^{p}\left(\frac{d_i}{z}\right)^2$ is called the severity or squared poverty gap index.
- $\boldsymbol{FGT_\alpha^k} = \frac{1}{N_k}\sum_{i=1}^{m_k}\left(\frac{d_i}{z}\right)^\alpha$ is the FGT index for subgroup k from a decomposed population, where $N_k$ is the total population and $m_k$ is the number of poor individuals in subgroup k.

Since in a poverty analysis, multiple indices have to be calculated concerning a given population to understand the level of poverty and well-being from various perspectives, in this work, we will compute the FGT index for values of α equal to 0, 1, and 2.

## 2.4 Logistic Regression

Since the poverty status in our context is a binary variable, 1 for poor and 0 for non-poor, we will use logistic regression to come up with a model best explaining and predicting the poverty status of a household based on their level of education.

### 2.4.1 Definition

Regression is one of the methods developed by Francis Galton that attempts to analyze the relationship between a dependent variable and one or more independent variables by changing one or multiple variables to understand the change happening in another (Hosmer, Lemeshow, and Sturdivant 2013). The dependent variables are called endogenous, while the independent ones are exogenous or predictive.

Logistic regression is a binomial model setting up a relationship between one or more explanatory variables and any type of dependent variable; it can be used in the following domains:

- **Marketing:** Either when there is a need to understand the impact of a marketing program or, conversely, why customers are switching companies.
- **Insurance:** Identifying clients likely to purchase an insurance policy against a particular risk.
- **Banking:** To assess risky loan applicants.
- **Medicine:** To determine the behaviors of sick patients compared to healthy subjects (Kleinbaum and Klein 2010).

### 2.4.2 Notation

Let $y$ be the qualitative variable to be predicted, and $x = (x_1, x_2, \ldots, x_n)$ the predictive variables. If $y$ is binary, meaning it only has two categories denoted as $\{1, 0\}$, and the $x_i$ are quantitative or qualitative, we refer to this as binary logistic regression.
We then have:

- Let $r$ be a set of nnn samples composed of $n_1$ and $n_{12}$ observations corresponding respectively to the categories 1 and 0 of $y$.
- $P(y = 1)$ and $P(y = 0)$ are the prior probabilities that $y$ is equal to 1 and 0, respectively. For simplicity, we will denote them as $P(1)$ and $P(0)$, respectively.
- $P(x|1)$ and $P(x|0)$ are the distributions of $\boldsymbol{x}$ given the value taken by $\boldsymbol{y}$, respectively.
- he posterior probabilities of obtaining category 1 and category 0 given xxx are denoted as $P(1|x)$ and $P(0|x)$, respectively.

### 2.4.3 Basic Hypothesis

Logistic regression is based on the following fundamental hypothesis called evidence:

$$\boldsymbol{EV(P) = ln\frac{p}{1-p}} \text{ with } \boldsymbol{p = P(1|x)}$$

Its logit model is then given by the following equation:

$$\boldsymbol{ln\frac{p}{1-p} = a_0 + \sum_{i=1}^{n} a_i\, x_i} \quad (\mathbf{1})$$

The ratio $\boldsymbol{\frac{p}{1-p}}$ is called the Odds.

Let's transform expression (1):

$$\text{De (1) on a } \frac{p}{1-p} = e^{a_0 + \sum_{i=1}^{n} a_i x_i}$$

$$\Rightarrow p = e^{a_0 + \sum_{i=1}^{n} a_i x_i} - p e^{a_0 + \sum_{i=1}^{n} a_i x_i}$$

$$\Rightarrow p\left(1 + e^{a_0 + \sum_{i=1}^{n} a_i x_i}\right) = e^{a_0 + \sum_{i=1}^{n} a_i x_i}$$

$$\boldsymbol{\Rightarrow p = \frac{e^{a_0 + \sum_{i=1}^{n} a_i x_i}}{1 + e^{a_0 + \sum_{i=1}^{n} a_i x_i}}} \quad (\mathbf{2})$$

If we have only one explanatory variable, equation (2) becomes:

$$\boldsymbol{p = \frac{e^{a_0 + a_1 x}}{1 + e^{a_0 + a_1 x}}} \quad (\mathbf{3})$$

If $\boldsymbol{p} > 0{,}5$, the individual is declared positive. Otherwise, the individual is declared negative.

**d. Estimation**

To construct the logit model based on a list of observations, one has to estimate the parameters a_i. The method of least squares can no longer be used here. For estimating the parameters, a technique known as the maximum likelihood method is used instead. The likelihood contribution or probability that an individual j belongs to a class is given by:

$$\boldsymbol{P(y_j = 1|x_j)^{y_j} \times (1 - p(y_j = 1|x_j))^{1-y_j}}$$

Therefore, the likelihood of any given sample is:

$$\boldsymbol{L = \prod_j P(y_j = 1|x_j)^{y_j} \times \left(1 - p(y_j = 1|x_j)\right)^{1-y_j}}$$

The most well-known and widely used optimization algorithm by software is the Newton-Raphson method (Hosmer, Lemeshow, and Sturdivant 2013).

**e. Overall Validity of the Model**

This step focuses on testing the significant role of an individual variable, or more specifically, how much a single explanatory variable contributes to the explanation of the dependent variable. If this contribution is significant, the variable is included in the equation; otherwise, it is excluded.

**f. Individual Evaluation of the Coefficients**

Finding the vector of parameters to estimate does not mean the problem is solved. Next, we need to verify the overall significance of the constructed model. To do this, we can use a test that is similar to the evaluation of a multiple linear regression model. We need to compare our model to a trivial model that is reduced to a constant only.

The hypotheses are as follows:

$$\begin{cases} H_0 : a_j = 0 \\ H_1 : a_j \neq 0 \end{cases}$$

This test is suited for the Wald statistic, which evaluates the null hypothesis that a given coefficient is equal to zero, meaning the variable does not contribute significantly to the model; if the test statistic is significant, the null hypothesis is rejected, meaning the variable should be included in the model.

Other evaluation procedures, such as the **Hosmer-Lemeshow** test, are not covered here.

.

### 2.5 The Confusion Matrix

#### 2.5.1 Definition

The confusion matrix is a contingency table used in classification problems to measure the quality of a prediction tool, sometimes called a classifier. In other words, it is a matrix that shows whether the prediction system or model succeeds in making correct predictions. It is obtained by cross-tabulating the number of observations of the estimated class in the columns and the number of occurrences of the actual class in the rows, containing reference data that must be different from the data used for the classification (Powers 2011).

#### 2.5.2 Actual or Reference Data

These are data collected in the field through surveys, aerial photographs, or thematic maps. However, they must be of the same typology as the classification data.

#### 2.5.3. Representation

Soit une variable explicative binaire ayant les modalités 1 pour positif et 0 pour négatif.

On a la matrice de confusion suivante :

Here is a typical layout for a confusion matrix:

Tableau 1: Représentation d'une matrice de confusion

| | **Classes prédites** |
|---|---|

| | | | |
|---|---|---|---|
| | | **1** | **0** |
| **Classes réelles** | **1** | **TP** | **FN** |
| | **0** | **FP** | **TN** |

- **True Positive (TP)**: The model correctly predicts the positive class.
- **False Negative (FN)**: The model incorrectly predicts the negative class when the actual class is positive.
- **False Positive (FP)**: The model incorrectly predicts the positive class when the actual class is negative.
- **True Negative (TN)**: The model correctly predicts the negative class.

## 3 Results

### 3.1 Evaluation of Household Expenditures for Education

#### 3.1.1 Education Expenditures in Kinshasa

Table 2: Education Expenditures per Household in Kinshasa

| | Non-Poor (CDF) | Poor (CDF) | Total (CDF) |
|---|---|---|---|
| Postgraduate | 350450 | 0 | 350450 |
| Primary | 2527950 | 1079070 | 3607020 |
| Non-Formal Program | 1383440 | 0 | 1383440 |
| Secondary | 12380730 | 2012988 | 14393718 |
| University | 20517820 | 419400 | 20937220 |
| None | 1407900 | 161450 | 1569350 |
| Total | **38568290** | **3672908** | **42241198** |

Source: Author's Calculation

Table 3: Education Expenditures in Kinshasa as a Percentage per Household

| | Non-Poor | Poor | **Total** |
|---|---|---|---|
| Postgraduate | 0,83% | 0,00% | 0,83% |
| Primary | 5,98% | 2,55% | 8,54% |
| Non-Formal Program | 3,28% | 0,00% | 3,28% |
| Secondary | 29,31% | 4,77% | 34,08% |
| University | 48,57% | 0,99% | 49,57% |
| None | 3,33% | 0,38% | 3,72% |
| **Total** | **91,30%** | **8,70%** | **100,00%** |

Source: Author's Calculation

#### 3.1.2 Education Expenditures in Lubumbashi

Table 4: Education Expenditures in Lubumbashi per Household

| | Non-Poor (CDF) | Poor (CDF) | Total (CDF) |
|---|---|---|---|
| Postgraduate | 239100 | 21300 | 260400 |
| Primary | 512850 | 205654 | 718504 |
| Non-Formal Program | 243795 | 12000 | 255795 |
| Secondary | 5109905 | 1332190 | 6442095 |
| University | 3522850 | 141150 | 3664000 |
| None | 148800 | 87689 | 236489 |
| **Total** | **9777300** | **1799983** | **11577283** |

Source: Author's Calculation

Table 5: Education Expenditures in Lubumbashi as a Percentage per Household

| | Non-Poor | Poor | **Total** |
|---|---|---|---|
| Postgraduate | 2,07% | 0,18% | 2,25% |
| Primary | 4,43% | 1,78% | 6,21% |
| Non-Formal Program | 2,11% | 0,10% | 2,21% |
| Secondary | 44,14% | 11,51% | 55,64% |
| Universitaire | 30,43% | 1,22% | 31,65% |
| None | 1,29% | 0,76% | 2,04% |
| **Total** | **84,45%** | **15,55%** | **100,00%** |

Source: Author's Calculation

### 3.1.3 The Share of Household Expenditures Devoted to Education is Therefore:

Table 6: Share of Household Expenditures Devoted to Education

| | **Kinshasa** | | **Lubumbashi** | |
|---|---|---|---|---|
| Poor | 3672908Fc | 2,104% | 9777300Fc | 1,719% |
| Non-Poor | 38568290Fc | 5,438% | 1799983Fc | 5,123% |
| **Total** | 42241198Fc | 7,542% | 11577283Fc | 6,842% |

Source: Author's Calculation

## 3.2 Some Education Indicators

Table 7: Education Indicators

| | **Kinshasa** | **Lubumbashi** |
|---|---|---|
| **Literacy** | | |
| Literate | 89,63% | 84,28% |
| Illiterate | 10,37% | 15,72% |
| **At least primary education level** | | |
| No | 6,33% | 9,66% |
| Yes | 93,67% | 90,34% |
| **Highest level of education** | | |
| Postgraduate | 0,21% | 1,70% |
| Primary | 17,74% | 25,38% |
| Non-formal program | 3,22% | 2,08% |
| Secondary | 51,56% | 51,89% |
| University | 20,95% | 9,28% |
| None | 6,33% | 9,66% |
| **Highest diploma obtained** | | |
| None | 12,14% | 25,76% |
| Others | 0,52% | 1,33% |
| Brevet CO | 13,69% | 37,69% |
| Certificat EP | 14,32% | 2,08% |
| D4, A3 | 7,99% | 14,39% |
| D6, A2 | 20,44% | 0,38% |
| Graduat | 11,41% | 3,60% |

| | | |
|---|---|---|
| Licence | 7,68% | 4,55% |
| ND | 7,05% | 9,66% |
| PP5 | 4,77% | 0,57% |
| **Reason for stopping studies** | | |
| Others | 4,56% | 4,36% |
| School failure | 1,35% | 1,14% |
| Schools too far | 0,52% | 1,70% |
| Studies completed | 35,58% | 18,94% |
| Pregnancy, marriage | 7,05% | 4,73% |
| Disability, illness | 1,45% | 1,52% |
| Financial inability of parents | 29,88% | 54,17% |
| Preference for an apprenticeship | 17,01% | 9,66% |
| Too young | 0,21% | 0,38% |
| None | 2,39% | 3,41% |
| **Lire écrire** | | |
| No | 23,44% | 40,53% |
| Yes | 76,56% | 59,47% |

The high literacy rates of household heads in the two cities are noted: 89.63 % for Kinshasa and 84,28% Lubumbashi as for the level of education higher longer works only 9.28% household head Lubumbashi studied at university, versus to Kinshasa, where this percentage is 20.95%. The share of household heads with completed secondary level is roughly similar in both cities at 52%.

At the level of diplomas, without diploma is to be noted for 25.76% and it is Diploma CO which arrives in head with a frequency of 38%, being but modestly represented inside our sample (108 on them). D6 and A2 diplomas are the most predominant in Kinshasa (20.44%) followed by Brevet CO diploma holders (13.69%) as well higher-level graduates with11,41%.

However, the school dropout rate for financial reasons remains concerning, reaching 30% in Kinshasa and 54% in Lubumbashi.

### 3.3 Measuring Poverty

#### 3.3.1 Data Extraction and Preprocessing

The necessary information (household consumption expenditures for poverty and educational variables) are in files 1 &3, respectively. Work1ow used several methods on the SPSS software to merge these data appropriately. Moreover, as only Kinshasa and Lubumbashi are considered in our study (more detailed geographical analysis would not be elaborated since limited to these two cities), this feature made us more accessible to extract needed data from the software. Ultimately, a total of 964 records were found for Kinshasa and 528 records in Lubumbashi.

#### 3.3.2 Well-being Indicator

Given the information available to us, the chosen well-being indicator for this work is household consumption expenditures.

### 3.3.3 Poverty line

According to previous studies conducted in 2005 by the World Bank, AFRISTAT, and the INS using data from the 1-2-3 survey and other sources, the monetary poverty threshold is **118,355.0068 Fc** per person in the country. The summary of this study is presented in the following table:

**Table 8: Poverty Threshold**

| Poverty Line | Urban Area | |
|---|---|---|
| | **FCFA/year/person** | **FC/year/person** |
| Food | 123070 | 95037,6843 |
| Non-food | 30195 | 23317,32249 |
| **Total** | 153265 | **118355,0068** |

Source: World Bank, AFRISTAT, INS 2005

To find the equivalent of this value for our statistical unit, it needs to be multiplied by the average household size in Congo. This results in 614,000 Fc per household.

### 3.3.4 Monetary Poverty Indices

We may apply the well-being indicator and the poverty threshold value to the database using SPSS to determine the poverty status variable now that we are aware of them. This variable uses total consumption expenditures to separate impoverished families from non-poor households. To be more precise, households will be categorized as non-poor if their total expenses exceed 614,000 Fc. All other households will be deemed poor.

Following this computation, we get the following outcomes: 42.53% of Kinshasa's 964 households, or 410 of them, are impoverished. 353 out of 528 homes, or 66.86% of the total, are impoverished in Lubumbashi. The values of the FGTo or H index for the cities of Kinshasa and Lubumbashi, respectively, are represented by these proportions.

The following table also presents the values of other poverty indices:

**Table 9: Incidence, Depth, and Severity of Poverty in Kinshasa and Lubumbashi**

| | $FGT_0$ | $FGT_1$ | $FGT_2$ |
|---|---|---|---|
| **Kinshasa** | **42,53%** | **13,04%** | **5,82%** |
| **Lubumbashi** | **66,86%** | **34,55%** | **20,69%** |

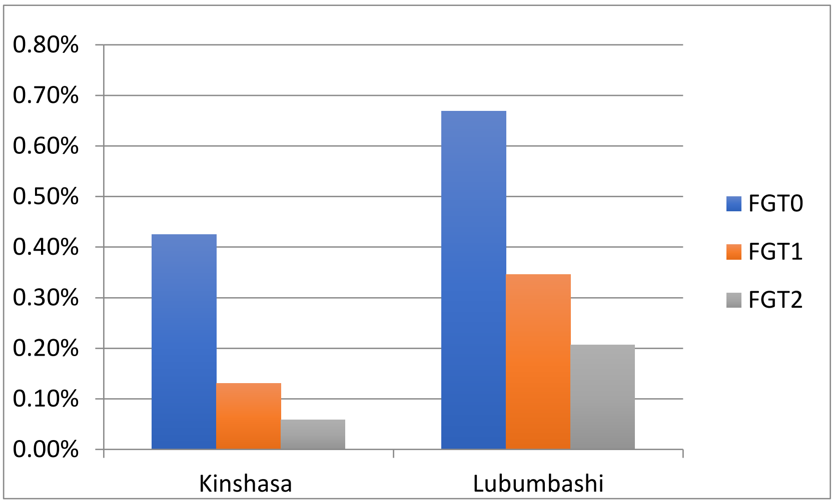
Source of the table: Author's Calculation

We notice that the poverty rate is higher in Lubumbashi than in Kinshasa: households in Kinshasa are less poor than those in Lubumbashi. This difference can be explained by the fact that Kinshasa, as the capital, hosts most of the country's major public and private institutions. Its airport, N'djili, is the main gateway in and out of the country. Additionally, being a very populous city, Kinshasa benefits from an abundant labor force. These factors give it significant advantages compared to Lubumbashi.

Regarding the depth and severity of poverty ($FGT_1$ and $FGT_2$ indices), the trend remains the same. The values of these indices for Lubumbashi are higher than those for Kinshasa, indicating that poverty is deeper and more severe there.

The following graph also summarizes this situation best:

Figure 1: Incidence, Depth, and Severity of Poverty in Kinshasa and Lubumbashi




• Poverty Profile by Education Level

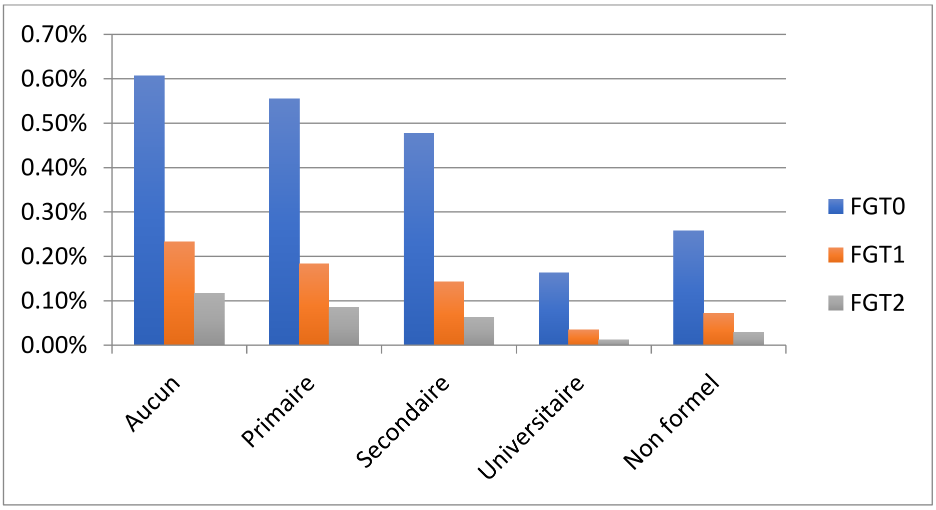
- **For the City of Kinshasa**

Table 10: Poverty and Education in Kinshasa

| Education Level | $FGT_0$ | $FGT_1$ | $FGT_2$ |
|---|---|---|---|
| **None** | 60,66% | 23,26% | 11,69% |
| **Primary** | 55,56% | 18,35% | 8,56% |
| **Secondary** | 47,69% | 14,29% | 6,24% |
| **University** | 16,34% | 3,42% | 1,21% |
| **Non-formal** | 25,81% | 7,14% | 2,86% |

This situation is graphically represented as follows:

Figure 2: Poverty Levels by Education of Household Heads in Kinshasa




A household whose head has attained the highest education level has only a 16.34% chance of being poor, whereas those whose head has not reached primary education have a 60.66% chance of being poor. Among individuals who have attained secondary education, 47.69% are poor.

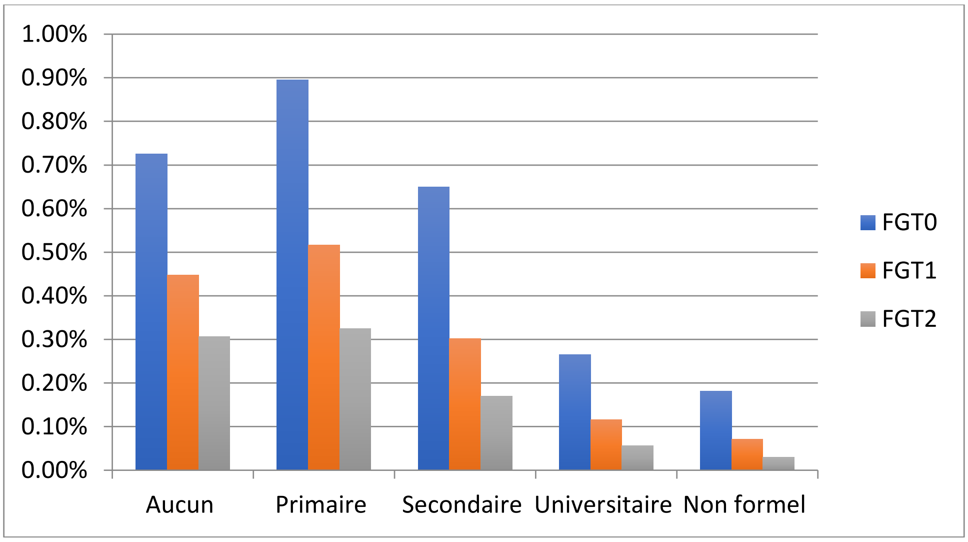
- **For the city of Lubumbashi**

Tableau 11: Poverty and Education in Lubumbashi

| Education Level | $FGT_0$ | $FGT_1$ | $FGT_2$ |
|---|---|---|---|
| **None** | 72,55% | 44,75% | 30,68% |
| **Primary** | 89,55% | 51,61% | 32,49% |
| **Secondary** | 64,96% | 30,23% | 16,96% |
| **University** | 26,53% | 11,61% | 5,63% |
| **Non-formal** | 18,18% | 7,11% | 2,95% |

This situation is represented graphically as follows :

Figure 3: Poverty Levels by Education of Household Heads in Lubumbashi




A household whose head has attained the highest education level has only a 26.53% chance of being poor, whereas one whose head has not reached primary education has a 72.55% chance. It's interesting to note that a household where the head has no education has better chances than one where the head has only reached primary education.

In summary, a household with a highly educated head has a lower likelihood of being poor, while those with heads who haven't attained the lowest education level have a much higher likelihood. This observation also holds for the depth and severity of poverty, indicating that the education level of household heads negatively influences household poverty status. Another observation is that poverty is widespread, affecting even those with secondary education levels, as will be confirmed in the next point.

## 3.4 Impact of Education on Poverty

### 3.4.1 Model Variables

The previous step allowed us to calculate the "poverty status" variable, which was used to measure poverty rates in the studied cities. Now, this variable will help establish the link between poverty and household education level. In other words, the "poverty status" variable will be used as the dependent or explained variable in our logistic regression model between poverty and education. For the explanatory or independent variables in the model, we will select some educational variables from the database based on their semantic relevance and significance. The selected variables are:

1. **Highest Degree Attained:**

- 0 = "None"
- 1 = "Certificat EP"
- 2 = "Brevet CO"
- 3 = "D4, A3"
- 4 = "PP5"
- 5 = "D6, A2"
- 6 = "Graduate"
- 7 = "Bachelor's Degree"
- 8 = "Doctorate"
- 9 = "Others"
- 10 = "Not Defined (ND)"

2. **Reasons for Stopping Education:**

- 1 = "Financial impossibility of parents"
- 2 = "Preference for an apprenticeship"
- 3 = "Pregnancy, marriage"
- 4 = "Disability, illness"
- 5 = "School failure"
- 6 = "Too young"
- 7 = "Schools too far away"
- 8 = "Studies completed"
- 9 = "Other reasons"

### 3.4.2 Estimation and validation of the model

The tables of results obtained with SPSS after applying logistic regression are:

- **For Kinshasa**

**Table 12: Model Estimation for Kinshasa**

**Variables in the Equation**

| | | B | S.E. | Wald | df | Sig. | Exp(B) |
|---|---|---|---|---|---|---|---|
| Step 1[a] | m13d_dip | | | 27,840 | 9 | ,001 | |
| | m13d_dip(1) | ,030 | ,324 | ,008 | 1 | ,927 | 1,030 |
| | m13d_dip(2) | ,000 | ,323 | ,000 | 1 | 1,000 | 1,000 |
| | m13d_dip(3) | -,215 | ,321 | ,449 | 1 | ,503 | ,807 |
| | m13d_dip(4) | -,395 | ,369 | 1,145 | 1 | ,285 | ,674 |
| | m13d_dip(5) | -,553 | ,418 | 1,747 | 1 | ,186 | ,575 |
| | m13d_dip(6) | -,338 | ,323 | 1,096 | 1 | ,295 | ,713 |
| | m13d_dip(7) | -1,185 | ,398 | 8,839 | 1 | ,003 | ,306 |
| | m13d_dip(8) | -1,994 | ,510 | 15,310 | 1 | ,000 | ,136 |
| | m13d_dip(9) | -1,127 | 1,165 | ,936 | 1 | ,333 | ,324 |
| | m17_are | | | 25,729 | 8 | ,001 | |
| | m17_are(1) | ,365 | ,339 | 1,155 | 1 | ,282 | 1,440 |
| | m17_are(2) | -,136 | ,354 | ,147 | 1 | ,701 | ,873 |
| | m17_are(3) | -,669 | ,401 | 2,784 | 1 | ,095 | ,512 |

| | | | | | | |
|---|---|---|---|---|---|---|
| m17_are(4) | ,297 | ,632 | ,221 | 1 | ,638 | 1,346 |
| m17_are(5) | ,801 | ,686 | 1,361 | 1 | ,243 | 2,227 |
| m17_are(6) | 21,013 | 28420,722 | ,000 | 1 | ,999 | 1335866972,656 |
| m17_are(7) | -,602 | ,965 | ,389 | 1 | ,533 | ,548 |
| m17_are(8) | -,449 | ,367 | 1,499 | 1 | ,221 | ,638 |
| Constant | ,190 | ,352 | ,292 | 1 | ,589 | 1,209 |

The model is valid at this stage since the Wald values for the parameters of two variables are significant.

**Table 13: Confusion Matrix for the Kinshasa Model**

| Confusion Matrix for the City of Kinshasa | | | | | |
|---|---|---|---|---|---|
| Observed | | | Predicted | | |
| | | | The household's monetary poverty status | | Percentage correct |
| | | | no poor | poor | |
| | The monetary poverty status of a household. | no poor | 372 | 166 | 69,1 |
| | | poor | 161 | 242 | 60,0 |
| | Overall percentage | | | | **65,2** |

Source: Author's calculation

This matrix shows that the prediction rate of the model for the city of Kinshasa is 65.2%. In other words, with this model, about 3 out of 5 households were correctly predicted. Although this is not highly significant, the model remains valid as its prediction rate already exceeds 50%.

- **For Lubumbashi**

Tableau 14: Model Estimation for Lubumbashi

Variables in the Equation

| | | B | S.E. | Wald | df | Sig. | Exp(B) |
|---|---|---|---|---|---|---|---|
| Step 1[a] | m13d_dip | | | 29,952 | 9 | ,000 | |
| | m13d_dip(1) | 1,117 | ,467 | 5,730 | 1 | ,017 | 3,057 |
| | m13d_dip(2) | ,049 | ,401 | ,015 | 1 | ,902 | 1,050 |
| | m13d_dip(3) | -1,227 | ,890 | 1,903 | 1 | ,168 | ,293 |
| | m13d_dip(4) | -,159 | ,758 | ,044 | 1 | ,834 | ,853 |
| | m13d_dip(5) | -21,746 | 21561,005 | ,000 | 1 | ,999 | ,000 |
| | m13d_dip(6) | -,795 | ,465 | 2,920 | 1 | ,087 | ,452 |
| | m13d_dip(7) | -,886 | ,670 | 1,750 | 1 | ,186 | ,412 |
| | m13d_dip(8) | -1,873 | ,768 | 5,948 | 1 | ,015 | ,154 |
| | m13d_dip(9) | 21,327 | 40192,970 | ,000 | 1 | 1,000 | 1828021964,901 |

| | | | | | | |
|---|---|---|---|---|---|---|
| m17_are | | | 30,807 | 8 | ,000 | |
| m17_are(1) | -,584 | ,657 | ,789 | 1 | ,374 | ,558 |
| m17_are(2) | -,867 | ,714 | 1,474 | 1 | ,225 | ,420 |
| m17_are(3) | -,720 | ,820 | ,771 | 1 | ,380 | ,487 |
| m17_are(4) | -2,124 | 1,011 | 4,415 | 1 | ,036 | ,120 |
| m17_are(5) | -3,100 | 1,152 | 7,240 | 1 | ,007 | ,045 |
| m17_are(6) | -1,790 | 1,556 | 1,323 | 1 | ,250 | ,167 |
| m17_are(7) | -1,652 | ,972 | 2,887 | 1 | ,089 | ,192 |
| m17_are(8) | -1,889 | ,696 | 7,365 | 1 | ,007 | ,151 |
| Constant | 1,765 | ,708 | 6,213 | 1 | ,013 | 5,843 |

The model is valid at this level because the Wald values for the parameters of two variables are significant.

Table 15: Confusion Matrix for the Lubumbashi Model

| **Confusion Matrix for the City of Lubumbashi** | | | | | |
|---|---|---|---|---|---|
| Observed | | | Predicted | | |
| | | | statut de pauvreté monétaire d'un ménage | | Percentage correct |
| | | | no poor | poor | |
| | The monetary poverty status of a household. | no poor | 81 | 84 | 49,1 |
| | | poor | 38 | 307 | 89,0 |
| | Overall percentage | | | | **76,1** |

**Source:** Author's calculation

Regarding Lubumbashi, the prediction rate is significant. It stands at 76.1%. This means that about 3 out of 4 households were correctly predicted by the model. Therefore, it is acceptable.

## 4 Interpretation and discussion

### 4.1 Estimation and Model Validation

The results obtained after applying logistic regression using SPSS for Kinshasa and Lubumbashi provided interesting insights into the relationship between poverty and education level.

#### 4.1.1 For Kinshasa

- Model Estimation The logistic regression model for Kinshasa shows that certain educational variables have a significant impact on poverty. The Wald values for these variables indicate the validity of the model.
- Confusion Matrix The confusion matrix for Kinshasa reveals an overall prediction rate of 65.2%, meaning the model correctly predicted the poverty status of 3 out of 5 households. While this rate is not extremely high, it surpasses 50%, indicating a useful model for predicting poverty in this city.

#### 4.1.2 For Lubumbashi

- Model Estimation The logistic regression model for Lubumbashi also indicates that certain educational variables significantly impact poverty, with significant Wald values for several parameters, confirming the model's validity.
- Confusion Matrix The confusion matrix for Lubumbashi shows an overall prediction rate of 76.1%. The model correctly predicted the poverty status of 76.1% of households, demonstrating its robustness for this city.

### 4.2 Comparative Analysis of Results

Comparing the results between Kinshasa and Lubumbashi reveals significant differences in poverty prediction. In Kinshasa, although the model is valid, it achieves a prediction rate of only 65.2%, while in Lubumbashi, the model achieves a rate of 76.1%. This difference could be due to contextual and socio-economic variations between the two cities, as well as differences in the quality and accessibility of education.

#### 4.2.1 Significant Variables

The results show that educational levels, particularly higher degrees, play a crucial role in poverty reduction. However, the impact of these degrees varies between the two cities. In Kinshasa, doctorate and bachelor's degrees have significant impacts, whereas in Lubumbashi, graduate, D4, and A3 degrees are more influential.

#### 4.2.2 Prediction Rates

The higher prediction rate in Lubumbashi may indicate better alignment of the model with the socio-economic reality of the city. The results underscore the importance of contextualizing education and poverty alleviation policies, taking into account local specificities.

#### 4.2.3 Highest Degree Attained

In Kinshasa, only individuals with a doctorate have a very low chance of being poor. Bachelor's degree holders also have a better chance of escaping poverty, while lower degrees have no significant impact. This reflects a highly competitive job market and a lower-quality education system that fails to create enough jobs. In Lubumbashi, a graduate, D4, or A3 degree helps reduce poverty, while a bachelor's degree has a lesser impact and a doctorate remains highly effective.

#### 4.2.4 Reasons for Discontinuing Education

Individuals who have completed their studies experience decreasing levels of poverty, especially in Lubumbashi. In Kinshasa, these individuals account for 35.58%, compared to 18.38% in Lubumbashi. This means that completing education increases the chances of success.

## 5 Conclusion and recommendations

The research successfully tackled each task as planned. To begin with we clarified the core concepts related to education and poverty assessing poverty using both monetary methods. The focus was, on poverty explaining the tools used for analysis such as the poverty threshold and common indicators.

After that we outlined the methodology primarily utilizing regression along with performance metrics for the model. This approach allowed us to explore the connection between a variable and multiple independent variables. We discussed the 2005 INS survey in detail, which supplied the data for our analysis (INS 2005).

For data examination, SPSS and Excel were employed. The theories on poverty helped us calculate the variable "poverty status," which served as the dependent variable in our logistic model. Measures of poverty ($H$ or $FGT_0$ or $FGT_1$ and $FGT_2$ indices) were computed, showing poverty rates of 42.53% in Kinshasa and 66.89% in

Lubumbashi. Additionally, we evaluated the depth and severity of poverty revealing percentages of 13.04% and 5.82% in Kinshasa and 34.55% and 20.69% in Lubumbashi respectively – emphasizing poverty conditions, in Lubumbashi.

We looked into education data. Found that many students drop out of school because of issues with rates, at 29.88% in Kinshasa and 54.17% in Lubumbashi. When we examined poverty based on the level of education of the household head, we observed that poverty tends to decrease as education levels increase but still persists. In Kinshasa 47.69% of individuals with an education are living in poverty while the figure is higher at 64.96%, in Lubumbashi.

Subsequently, logistic regression was applied between "poverty status" and educational variables "highest degree obtained" and "reasons for discontinuing education." Results indicated that a doctorate offers the best chance to escape poverty. In Kinshasa, bachelor's degree holders also have some chance of success, while in Lubumbashi, graduate, D4, and A3 degrees play crucial roles.

This study illustrates widespread poverty, especially in Kinshasa, where even those with secondary education are not exempt. In the DRC, despite education, escaping poverty is not guaranteed, raising questions about the quality of education provided.

**Recommendations:**

- Strengthen Education Policies: The government, in collaboration with development agencies, should enhance the quality of education at all levels, making education more accessible and addressing issues like hunger that hinder learning.
- Combat Negative Social Norms: Phenomena like "transactional sexual relations" and "commodification of education" must be eradicated to ensure quality education.
- Facilitate Access to Scholarships: Increase local and international scholarship opportunities to motivate students.
- Vocational Training Programs: Institute programs to enhance skills among university graduates and reduce unemployment.
- Develop Higher Education: Build higher education institutions and promote postgraduate studies to increase the number of professors and improve learning conditions.
- Encourage Entrepreneurship: Support youth entrepreneurship as a means to combat unemployment.
- Internal Migration: Encourage graduates to migrate within the country to increase human capital and reduce excessive labor market competitiveness in the capital.

These measures will enhance education quality and alleviate poverty in the DRC, particularly in Kinshasa and Lubumbashi.

**REFERENCES**

[1] Afristat. 1997. 1-2-3 Survey Methodology. Accessed July 02, 2024. https://www.afristat.org.

[2] Banerjee, Abhijit V., and Esther Duflo. 2011. Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty. New York: PublicAffairs.

[3] Becker, Gary S. 1993. Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. 3rd ed. Chicago: University of Chicago Press.

[4] Glewwe, Paul, and Michael Kremer. 2006. "Schools, Teachers, and Education Outcomes in Developing Countries." Handbook of the Economics of Education 2: 945-1017.

[5] Hanushek, Eric A., and Ludger Woessmann. 2008. "The Role of Cognitive Skills in Economic Development." Journal of Economic Literature 46 (3): 607-668. https://doi.org/10.1257/jel.46.3.607.

[6] Hosmer, David W., Stanley Lemeshow, and Rodney X. Sturdivant. 2013. Applied Logistic Regression. 3rd ed. Hoboken, NJ: Wiley.

[7] International Monetary Fund (IMF). 2015. Democratic Republic of Congo Poverty Report. October.

[8] King, Elizabeth M., and Rebecca Winthrop. 2015. "Today's Challenges for Girls' Education." Global Economy and Development at Brookings Working Paper 90. https://www.brookings.edu/research/todays-challenges-for-girls-education.

[9] Kleinbaum, David G., and Mitchel Klein. 2010. Logistic Regression: A Self-Learning Text. 3rd ed. New York: Springer.

[10] National Institute of Statistics (INS). 2005. 1-2-3 Survey on Employment and the Informal Sector. Democratic Republic of Congo.

[11] Psacharopoulos, George, and Harry Anthony Patrinos. 2004. "Returns to Investment in Education: A Further Update." Education Economics 12 (2): 111-134. https://doi.org/10.1080/0964529042000239140.

[12] Sachs, Jeffrey D. 2005. The End of Poverty: Economic Possibilities for Our Time. New York: Penguin Press.

[13] Schultz, Theodore W. 1961. "Investment in Human Capital." The American Economic Review 51 (1): 1-17.

[14] Sen, Amartya. 1999. Development as Freedom. New York: Knopf.

[15] Todaro, Michael P., and Stephen C. Smith. 2015. Economic Development. 12th ed. Boston: Pearson.

[16] UNESCO. 2023. "Education Transforms Lives." Accessed July 04, 2024. https://en.unesco.org/themes/education.

[17] United Nations. 2023. "Sustainable Development Goal 4: Quality Education." Accessed June 18, 2024. https://sdgs.un.org/goals/goal4.

[18] World Bank. 2023. "Education." Accessed July 18, 2024. https://www.worldbank.org/en/topic/education.

[19] Mandela, Nelson. 2003. *Lighting Your Way to a Better Future: Speech delivered by Mr. N. R. Mandela at the launch of Mindset Network.* Johannesburg, South Africa, July 16, 2003.

[20] Foster, James, Joel Greer, and Erik Thorbecke. 1984. "A Class of Decomposable Poverty Measures." *Econometrica* 52 (3).

[21] Peng, Chao-Ying Joanne, Kuk Lida Lee, and Gary M. Ingersoll. 2002. "An Introduction to Logistic Regression Analysis and Reporting." *The Journal of Educational Research* 96 (1).