



## THE CHARACTERISTICS AND SPATIAL MODELLING OF INFORMAL ECONOMIES



**Dounya Matsop  
Claude<sup>1+</sup>  
Ngouhouo Ibrahim<sup>2</sup>  
Joseph Emmanuel  
Fantcho<sup>3</sup>  
Tchuani Ngakam  
Ulrich Ndesil<sup>4</sup>**

<sup>1,2</sup>University of Dschang, Cameroon.

<sup>1</sup>Email: [claudeanal2008@yahoo.fr](mailto:claudeanal2008@yahoo.fr) Tel: (+237)675966360

<sup>4</sup>Email: [ulrichndesil@gmail.com](mailto:ulrichndesil@gmail.com) Tel: (+237)697792738

<sup>2</sup>Professor, University of Dschang, Cameroon.

<sup>2</sup>Email: [ngouhouo@yahoo.fr](mailto:ngouhouo@yahoo.fr) Tel: (+237)696662260

<sup>3</sup>Universite des Montagnes, Bagangte, Cameroon.

<sup>3</sup>Email: [je.fantcho@gmail.com](mailto:je.fantcho@gmail.com) Tel: (+237)699309155



(+ Corresponding author)

### ABSTRACT

#### Article History

Received: 14 April 2020

Revised: 18 May 2020

Accepted: 22 June 2020

Published: 9 July 2020

#### Keywords

Informal economy

Data analysis

Aggregate index

Logistic regression.

#### JEL Classification:

E26 Informal Economy;

Underground Economy

The goal of this study is to identify the characteristics of informal economies. Specifically, we identify the factors that describe informal economies and summarize them into indices measuring informality. We use data on 189 countries and the method used to analyze the data focuses only on the year 2012. We perform an exploratory analysis to identify the variables which structure informal economies and use the scores from a logistic regression to measure the degree of informality of each country. The results show that the degree of informality a country is related to its level of development. Thus, developing countries are generally those where the degree of informality is highest while developed countries are generally characterized by a low level of informality. This study thus enables us to classify countries into groups according to the variables that determine informality and draw a chart representing the countries according to their level of informality.

**Contribution/ Originality:** This study contributes to the existing literature by identifying the characteristics of informal economies.

## 1. INTRODUCTION

In the majority of African and Latin American countries, Structural Adjustment Programs (SAPs) led to an important reduction of government spending, a progressive stop of investment projects and a rationing of the personnel in the public service (Walther, 2006). Adjustment policies deeply affected the structure of employment, forcing the highly indebted poor countries to go from an organization where the State is the main employer to one where individual and private initiative must take over.

The importance of the informal economy is a growing in certain countries. In Cameroon for example, the resources of the informal sector accounted for 90,4% of the total resources in 2005; this number increased to 90,5% in 2010<sup>1</sup>.

Schneider (2004) measures the share of the value-added of the informal economy in the GDP<sup>2</sup> (Gross domestic product) of 145 countries across the world. According to the author, the informal economy averagely accounts for

<sup>1</sup> Cameroon Employment and informal sector survey (2005) and (2010).

43,2% of the GDP of African economies; 30,8% of the GDP of Asian countries; 43,4% of the GDP of Central and South America; 40,1% of the GDP of Central and Eastern Europe; and 16,3% of the GDP of OCED countries in 2003. The author combines the Multiple Indicators, Multiple Causes (*MIMIC*) method and monetary approach to measure the size of the informal economy between 1999 and 2003.

The informal economy refers to the set of economic activities carried out by people not registered with the social security<sup>2</sup>. However, the structure of the informal economy depends on the level of development of the country. In developing countries, the informal sector generates the majority of employment in cities, the shantytowns and villages. It offers less expensive goods and services adapted to local realities and the populations needs.

The informal economy includes petty-trading, craftsmen and family small-scale businesses which barely survive in their activities. Informal jobs in developing countries resemble a struggle for survival. For this reason, informal activity refers to the exercise by the poor of jobs that require painful tasks that are not recognized, recorded, protected, or regulated by the public authorities. The role of the State is thus called into question in the development of the informal economy. The deployment of the informal economy is a strategy of survival of the poor and vulnerable. According to [Maldonado \(2001\)](#) informal activity is the only alternative offered to the unemployed and newcomers in the job market in Africa.

On the other hand, in the developed countries the State occupies a dominant position in economic activity. The informal economy evolves according to rules and regulations imposed by the State. The size of the informal economy depends on the level of taxes and other levies deducted from exchange.

According to [De Soto \(1989\)](#) the informal sector is made up of agents who choose to act in an informal manner to avoid the costs, time and effort required to register their activity. The analysis of De Soto is valid in developed countries where informal work, also referred to as moonlighting is a risky activity and severely punished by the tax authorities.

According to [Maloney \(2004\)](#) the informal sector is consists of the agents who deliberately seek to flee payments, levies, and other costs associated with economic activity to seek illegal goods and services.

In a general manner, the informal economy has two components: the first refers to illegal and criminal activities (prostitution, drugs) and the second component refers to the production of legal but unregistered goods and services.

This study makes a review of the factors which characterize informal economies and summarizes them into indices of measurement of the level of informality in each country. While, the second part will focus on the review of the literature, sections 3 and 4 will successively present the methodology and the results. As for the last section, it will conclude the work.

## 2. REVIEW OF LITERATURE

### 2.1. Theoretical Approaches

In the economic literature, three dominant approaches are used to understand the origins and causes of informality ([Bacchetta, Ernst, & Bustamante, 2009](#)):

- The dualistic approach: This approach falls under the extensions of [Lewis \(1954\)](#) and [Harris and Todaro \(1970\)](#). It is based on a dual model of the job market where the informal sector is regarded as a residual component of this market that is not related to the formal economy. It is a subsistence economy that exists only because the formal economy is unable to offer a sufficient number of jobs.

---

<sup>2</sup> Friedrich Schneider use the term "shadow economy".

<sup>3</sup> This definition pre-supposes that there exists a national social security fund that registers all workers officially exercising an economic activity.

- The structural approach: It highlights the interdependences between the informal and formal sectors. According to this approach Marxist oriented approach, the informal sector is integrated in the capitalist system following a subordination relationship by providing labor and goods at a cheap rate to formal companies.
- The legal approach: It considers that the informal sector is made up micro-entrepreneurs who prefer to operate in an informal way to escape public regulations considered too tough. It is different from the two former approaches in that the choice of informality is voluntary and related to the excessive costs of legalization associated with the formal statute and registration.

New methods of measurement make it possible to quantify the respective weights of these sectors and analyze the production dynamics.

## 2.2. Empirical Approach

### 2.2.1 Measuring the Level of the Informality: Which Approach is Adequate?

To measure the size of the informal economy, of many methods and variables are used. Some of these methods include:

- Direct or microeconomic methods: These methods are based on the households and company surveys as well as tax audits. These methods specifically compare household expenditure to the disposable income of the households; wages paid to the taxes levied on these wages; the sale of goods and services subjected to the VAT (Value Added Tax) to the amount of VAT received; production against the tax on production; data from expenditure surveys against tax data tax or data on incomes; data from company surveys against tax data; the real value of VAT against theoretical VAT; real income tax returns against theoretical income tax; uses of labor against labor resources.
- The direct approaches have the advantage of providing detailed information on the structure of the informal economy and the disadvantage of not being able to estimate the development and the growth of the informal economy over a long period.
- Indirect or macroeconomic methods: These approaches use many indicators likely to provide information on the informal economy in time. We distinguish:
- The approach through the GDP: This approach confronts the GDP calculated through the expenditure approach to the GDP calculated following the income approach. According to this approach:

$$\text{Size of the informal economy} = \frac{(\text{GDP expenditure approach}) - (\text{GDP income approach})}{(\text{GDP income approach})}$$

- The employment approach: This is based on the indicators of the labor market and supposes that the fall in the rate of employment in the formal sector of the economy should be perceived as an indicator of an increase in informal activities (Schneider, 2000). It properly does not measure the size of the informal economy.
- Monetary methods<sup>4</sup>: Cagan (1958) was the first to use the approach through the demand for money to estimate the size of the informal economy. According to him, the tax pressure is one of main determinants of the size of the informal economy. The result of his studies carried out on the US economy (1958) show a correlation between the demand for money and the tax pressure in the United States over the 1919-1955 period. The author thus concludes that there exists a relationship between the demand for money and the size of the informal economy.

---

<sup>4</sup> The transactions approach (Feige, 1979) the method of the ratio of liquid assets (Gutmann, 1977) and the method of monetary demand (Tanzi, 1982).

- The transactions method: Feige (1979) is the first to use this approach. He goes from the following hypothesis: There exists a constant relationship between monetary transactions and the GDP in time. However, this hypothesis is not easily verifiable.
- The method of estimation by a global indicator proposed by Kaufmann and Kaliberda (1996): This method supposes the existence of a precise and stable relationship between the consumption of electricity and output. According to the authors, the elasticity of the consumption of electricity relative to aggregate output is very close to 1. This approach poses some concerns since in agriculture based economies, the described relationship is not stable. Also, in many developing countries, electricity is not the main source of energy. It should be noted that the consumption of electricity in each country also depends on its cost.
- Method of the latent variable developed by Frey and Weck (1983): This method analyzes the relationship between an unobserved dependent variable and several observed explanatory variables. According to this method, the size of the underground economy can be explained by the effective tax burden, the perception of the tax burden, the unemployment rate, the regulatory burden, the attitude towards the payment of taxes (tax morality) and the available income per capita.

The method of latent variables however has some limitations. Firstly, the concept of “tax morality” is difficult to measure in an objective manner. In addition, the results are not very reliable and are unstable. Helberger and Knepel (1988) show that a slight change in the countries used in the example of Frey and Weck leads to very different results. The ambiguity of the data and instability of the results limit the use of the method of latent variables as means of measurement of the underground economy.

### 2.2.2. The Informal Economy: Residue of the Formal Economy?

The “dualistic” approach of Harris and Todaro (1970) comes to support the idea we have of the shadow economy in developing countries. In fact, the dual model of the job market considers the informal sector as a residual component disconnected from the rest of the economy. It is thus a sector of subsistence whose existence results from the inability of the formal sector to create sufficient jobs.

According to Maurizio (2012) it is possible to establish a relationship between informality and poverty in Latin America. Muheme (1995) holds that the informal economy is a form of disguised unemployment.

The growth of the informal economy is a survival strategy of the poor and the vulnerable. For Maldonado (2001) informal activity is the only alternative offered to the unemployed and new arrivals on the job market in Africa.

De Soto (1989) poses the basis for the analysis of the informal economy. His analysis is rather valid in developed countries where informal activity, also referred to as moonlighting is a risky activity severely punished by the tax department.

Also, the results of Lacroix and Fortin support the ideas of the structural doctrine (Portes, Castells, & Benton, 1989). In fact, the structural approach focuses on the interactions between the informal and formal sectors. The informal sector is a sub-component of the capitalist system and provides the formal companies cheap goods and labor.

Tanzi (1980) proposes an econometric approach to measure the size of the informal economy. The econometric method introduces the tax burden into the equation of the demand for money. Tanzi specifically seeks to isolate the excess demand for currency due to informal transactions. Contini (1981) uses the approach through the rate of participation in the job market to measure the size of the informal economy. This approach uses statistics on the job market to estimate the size of the informal economy. Kaufmann and Kaliberda (1996) use electricity consumption as a physical indicator of the global economic activity.

According to Frey and Weck (1983) the size of the underground economy can be explained by the effective tax burden, the perception of the tax burden, the unemployment rate, the level of regulation (for example: the number of laws), the attitude with regards to the payment of taxes (tax morality) and the income per capita.

The shadow economy functions partly thanks to corruption and has a cost for the agents. These agents must corrupt tax officials or face penalties and pay fines. De Soto (1989) finds that between 10 and 15% of the income of informal companies is paid in the form of bribe as against only 1% for their counterparts of the formal sector. The level of corruption and the burden of the regulation are thus variables which characterize informal economies<sup>5</sup>.

### 3. METHODOLOGY AND RESULTS

The objective of this section is to characterize the informal economy by presenting a theoretical model, which will later be verified using analytical data

#### 3.1. Presentation of the Model

##### 3.1.1. The Linear Model with Latent Variable

Latent variables are a first attempt to address the problems involved in the use of ordinary least squares in a model whose dependent variable is binary. The latent variable is a continuous non-observable variable which is representative of the studied phenomenon (for example the level or degree of informality of a country can be studied using the fact that certain indicators are related to the risk of informality). Thus, we introduce the latent variable  $y_i^*$  (income) and suppose that:

- the degree of informality of a country  $i$  is greater than the 0, then the country is informal, that is,  $Y_i = 1$  if

$$Y_i^* > 0.$$

- $Y_i^*$  is a linear function of  $X_i$  such that:  $Y_i^* = a_0 + a_1X_i + \varepsilon_i$  (1)

The binary dependent variable  $Y_i$  is then given by the following decision model:

$$Y_i = \begin{cases} 1 & \text{si } Y_i^* > 0 \\ 0 & \text{si } Y_i^* \leq 0 \end{cases} \quad (2)$$

This decision rule simply consists in supposing that the proportion of the variables ( $Y_i = 1$ ) is high for:

$$Y_i^* = a_0 + a_1X_i + \varepsilon_i > 0. \quad (3)$$

Consider  $P_i$  the probability that  $Y_i^* > 0$ . This probability is given by:

$$P_i = \text{Prob}(Y_i = 1) = \text{Prob}(Y_i^* > 0) = \text{Prob}(a_0 + a_1X_i + \varepsilon_i > 0) = \text{Prob}(\varepsilon_i > -(a_0 + a_1X_i)) \quad (4)$$

If the distribution of the error  $\varepsilon_i$  is centered relative to the mean, we obtain:

$$\text{Prob}(\varepsilon_i > -(a_0 + a_1X_i)) = \text{Prob}(\varepsilon_i < a_0 + a_1X_i) \quad (5)$$

<sup>5</sup> See table below which gives the list of variables retained for the exploratory analysis.

Thus: 
$$P_i = \text{Prob}(Y_i = 1) = \text{Prob}(\varepsilon_i < a_0 + a_1 X_i) \quad (6)$$

The proportion  $P_i$  thus depends on the distribution of the error term  $\varepsilon_i$  of the decision model.

In fact, in a probit model, the cumulative function of the error term  $\varepsilon_i$  is given by:

$$P_i = \int_{-\infty}^{(a_0 + a_1 X_i)} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \quad (7)$$

In addition, the logit or logistic function is given by:

$$\begin{aligned} P_i = \text{Prob}(Y_i = 1) &= \text{Prob}(Y_i^* > 0) = \text{Prob}(a_0 + a_1 X_i + \varepsilon_i > 0) = \text{Prob}(\varepsilon_i > -(a_0 + a_1 X_i)) \\ &= \Phi(a_0 + a_1 X_i) \end{aligned} \quad (8)$$

where  $\Phi(\cdot)$  is the cumulative distribution of the logistic function.

$$P_i = \Phi(a_0 + a_1 X_i) = \frac{\exp(a_0 + a_1 X_i)}{1 + \exp(a_0 + a_1 X_i)} = \frac{1}{1 + \exp(-(a_0 + a_1 X_i))} \quad (9)$$

The properties of this equation are as follows: if  $a_1 > 0$ , then  $\lim_{X_i \rightarrow +\infty} P_i \rightarrow 1$  and  $\lim_{X_i \rightarrow -\infty} P_i \rightarrow 0$  which

ensures that the values of  $P_i$  lie between 0 and 1. After being transformed, the Logit function can be written:

$$\text{Ln}\left(\frac{P_i}{1-P_i}\right) = Y_i^* = a_0 + a_1 X_i + \varepsilon_i \quad (10)$$

In order to determine the synthetic index, we model the degree of informality of the economy in country  $i$

Here, we explain the level of informality of the economy in each country taking into consideration certain indicators describing the risk of informality. We model the degree of informality of the economy  $Y_i^*$  in the country  $i$  in the following manner:

$$Y_i^* = X_i \beta + \sigma e_i \quad (11)$$

We choose the elements of  $X_i$  as being the basic indicators of informality. We suppose that the errors  $e_i$  are independent, of zero mean, and constant variance;  $\sigma^2 > 0$ . These error terms are independent of the explanatory variables. Under these assumptions, the expected level of informality that we also call the score is equal to:  $E(Y_i^*) = X_i \beta$ . If all the components of the vector  $\beta$  are positive, then the level of informality will be an increasing function of each component of  $X_i$ . The idea is that if the level of informality of the country  $i$  is higher

than a threshold  $c$ , then this country is informal. If  $Y_i$  is a random variable defining the a priori level of informality in the economy of country  $i$ .

Equation 11 can also be written:

$$Y_i^* = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} + \sigma e_i \quad (12)$$

$X_1$  : GDP PER CAPITAL.

$X_2$  : ELECTRIC POWER PER CAPIT.

$X_3$  : MONEY (M2) PER CAPITAL.

$X_4$  : AGRICULTURE VALUE ADDED PER.

$X_5$  : DOMESTIC CREDIT FINANCIAL.

$X_6$  : REPEATERS PRIMARY SCHOOL.

$X_6$  : TOTAL TAX RATE.

$X_8$  : COST BUSINESS START UP P.

$X_9$  : UNEMPLOYMENT OF YOUTH.

$X_{10}$  : AGE DEPENDENCY RATIO.

$\beta_0 = C$ : constancy or informality threshold.

However,  $Y_i$  is a binary variable given by:

$$Y_i = \begin{cases} 1 & \text{si } Y_i^* > c \\ 0 & \text{si } Y_i^* \leq c \end{cases} \quad (13)$$

Thus, we can calculate the probability<sup>6</sup> that country  $i$  is informal:

$$E(Y_i | X_i = x_i) = Prob(Y_i = 1 | X_i = x_i) = Prob(Y_i^* > c | X_i = x_i) = Prob\left(\frac{e_i}{\sigma} > \frac{c}{\sigma} - \frac{1}{\sigma} x_i \beta\right) \quad (14)$$

---

<sup>6</sup> This probability which we also call risk or level of informality is our aggregate informality index.



By supposing that the error term  $e_i$  in the expression (11) follows a logistic distribution<sup>7</sup>, we can calculate the likelihood function of the model given by:

$$p_i = Prob(Y_i = 1|X_i = x_i) = \frac{\exp(-\frac{c}{\sigma} + \frac{1}{\sigma}x_i\beta)}{1 + \exp(-\frac{c}{\sigma} + \frac{1}{\sigma}x_i\beta)} \tag{15}$$

Conditionally to the observed values of the explanatory variables, the likelihood function of the model is written:

$$l(y_1, \dots, y_N | x_1, \dots, x_N; c, \beta, \sigma) = \prod_{i=1}^N p_i^{y_i} (1 - p_i)^{1-y_i} \tag{16}$$

Where  $y_i \in \{0,1\}$   $i = 1, \dots, N$ ,  $N$  indicates the number of countries studied. By normalizing<sup>8</sup> with  $\sigma = 1$  using the method of maximum likelihood to identify the set  $\theta = \{c, \beta\}$  of parameters of the model.

We write the log-likelihood of the model as:  $logl(y|x; \theta) = \{\sum_{i=1}^N y_i \ln(p_i) + (1 - y_i) \ln(1 - p_i)\}$

The maximum likelihood estimator  $\hat{\theta}$  is given by:  $\hat{\theta} = \underset{\theta}{\operatorname{argmax}} \log l(y|x; \theta)$

Under the assumptions of regularity (Davidson & MacKinnon, 1993),  $\hat{\theta}$  is asymptotically normal and converges towards the true value of the vector of parameters  $\theta_0$ .  $\hat{\theta}$  is obtained by solving the following system of  $K+1$  equations:

$$\frac{\partial \log l(y|x; \hat{\theta})}{\partial \hat{\theta}} = \sum_{j=1}^N \frac{(y_i - F(x_i\beta - c))f(x_i\beta - c)}{F(x_i\beta - c)(1 - F(x_i\beta - c))} x_i = 0 \tag{17}$$

$f$  is the density of the logistic distribution, and  $F$  the cumulative distribution. For the logistic distribution, Equation 17 is simplified and can be written:

$$\frac{\partial \log l(y|x; \theta)}{\partial \theta} = \sum_{j=1}^N (y_i - F(x_i\beta - c)) x_i = 0 \tag{18}$$

In fact, by using the properties of the logit models, we can calculate:

$$odd(i) = \frac{p_i}{1-p_i} = \exp(-c + x_i\beta) \tag{19}$$

<sup>7</sup> $e_i$  Following a logistic distribution if the cumulative distribution  $F$  is:  $F(x) = Prob(e_i \leq x) = \frac{1}{1 + \exp(-x)}$ . The density function of the logistic

distribution is:  $f(x) = \frac{\exp(x)}{[1 + \exp(x)]^2}$ . Note that:  $Prob(e_i > -x) = \frac{\exp(x)}{1 + \exp(x)}$

<sup>8</sup> We normalise because the parameters  $c$  and  $\beta$  that in the likelihood are all divided by  $\sigma$  (see the expression of  $p_i$ ).



Odd ( $i$ ) is the odd related to the level of risk of informality of country  $i$ .  $p_i$  is the probability that  $Y_i = 1$  or the probability that country  $i$  is informal.  $1 - p_i$  is the probability that country  $i$  is not informal. If we estimate the parameters of the model and find for example that:  $\exp(-\hat{c} + x_i\hat{\beta}) = 4$ , this means that the probability that country  $i$  is informal is four times higher than the probability that this country is not informal.

Consider  $x_1^{[k]}$ , the  $K^{\text{th}}$  component of the vector of the vector  $x_i$ . We can calculate:

$$\frac{\partial p_i}{\partial x_1^{[k]}} = f(-c + x_i\beta) \cdot \beta_k \quad (20)$$

$\frac{\partial p_i}{\partial x_1^{[k]}}$  measures the rate of change of the degree of informality, following a marginal variation in the  $k^{\text{th}}$  criterion of informality. We can also calculate “the elasticity of informality”. Elasticity measures the proportion in which the probability (or degree) of informality varies following a 1% change in an indicator of informality (all things being equal). The elasticity of the risk of informality is given by:

$$\frac{\partial p_i}{\partial x_1^{[k]}} \cdot \frac{x_1^{[k]}}{p_i} = (1 - p_i)\beta_k x_1^{[k]} \quad (21)$$

### 3.1.2. Statistical Tests and Validation of Hypotheses

The estimation of the parameters is carried out here using the algorithms of maximum of a log-likelihood function (Thomas, 2002a). Only the signs of the coefficients indicate if the variable acts positively or negatively on the probability  $P_i$ . It is however possible to calculate the marginal effects (Thomas, 2002b) in order to know the sensitivity of changes in an explanatory variable on the probability  $P_i$ . The significance of the coefficient is thus appreciated using the ratios known as “Z-Statistics” because the distribution of the ratios of the coefficient on its standard deviation does not follow a Student law as in the general linear model, but a normal distribution.

In order to test the hypothesis  $H_0: a_1 = a_2 = a_3 = \dots = a_k = 0$ , we use the log-likelihood ratio statistic given by:

$$LR = -2(\ln(L_R) - \ln(L_U)) \quad (22)$$

Where  $L_R$  represents the value of the log-likelihood function under the  $H_0$  hypothesis and  $L_U$  the value of the unconstrained log-likelihood function. Under the null hypothesis  $H_0$ ,  $LR$  follows a Chi-square ( $\chi^2$ ) distribution with  $k$  degrees of freedom.

Decision rule: If the  $LR$  statistic is higher than the  $\chi^2$  value read on the table for the given level of significance (generally 5%), then we reject the null hypothesis  $H_0$ . Thus, the estimated model has at least one significant explanatory variable. Given the characteristic of the dependent variable coded 0 or 1, the coefficient of determination  $R^2$  is not interpretable in terms of adjustment of the model. This is why we use the statistic called “the pseudo- $R^2$ ” which is given by:

$$R^2 = 1 - \frac{\text{Log}(Lu)}{\text{Log}(L_R)} \quad (23)$$

**Table-1.** Variables retained for an exploratory analysis of informality.

Nature of the Variables	Variables selected	Approach
Physical variable Electricity	Consumption of electricity in kWh per capita	Kaufmann and Kaliberda*
	Proportion of the population with access to electricity	
Monetary variables	Money supply expressed as a percentage of the GDP	Monetary approach*
	Money and quasi money (M2) per capita	
Income per head	GNI per capita	Approach by the GDP*
	GDP per capita	
Capital per head	Gross fixed capital formation per capita ( current US \$)	Variables judged important
Savings per head	Gross domestic savings per head	
Expenditure per head	Final consumption expenditure of households per capita	Approach: Direct*, Monetary*, GDP*
	Government final consumption expenditure	
	Health expenditure per capita (current US \$)	
gender	Proportion of seats occupied by women in the national parliament	Taking into account of the sex
Education	Expenditure on education per capita (US \$)	Role of education, ILO
	Rate of failures in primary education	
	Rate of failure in secondary education	
Agriculture	Added-value by worker of Agriculture (US \$)	Residual approach*
Unemployment	Unemployment rate	Approaches by employment
	Youth unemployment rate	
	Dependency ratio (% Of the working-age population)	
Burden of the regulation	number of hours to prepare and pay the taxes	De Soto*, method of latent variable*
	Number of days necessary to enforce a contract	
	Cost of procedures of starting of a company	
	Quality of the regulation and business environment	
	Number of day necessary to start a company	
	Number of procedures to record a new company	
Taxation	Total tax rate expressed as a percentage of the trading profits	De Soto*, MIMIC* (latent variable)
	Consolidated rate of customs tariff (%)	
Banking environment	Number of adults (on 1 000) that deposit income in banks	Monetary approaches*
	Share of domestic credit provided by the financial sector (% GDP)	

### 3.2. Variables and Nature of the Data

#### 3.2.1. Variables

We retain some variables that can explain the levels of informality of economies. Table 1 summarizes the variables retained for an exploratory analysis of the dimensions of informality. The variables are selected following the literature on the measurement of the informal economy and depending on the availability of data on the variables. In this table, certain approaches or names of authors are marked by a star because of the existing difference between the variables selected and those presented in the literature.

We add some variables on education because according to the ILO, an easy means of integrating an official employment and have access to a decent job is the elimination of basic illiteracy. Social inequalities are the reflection of differences related to education. It is observed that the illiterate have no other choice than working in the informal economy. Thus, basic education is a good means of entering the formal sector. In fact, 40% of adults in sub-Saharan Africa are illiterate. This is also the case of about half of the adult population in South Asia (ILO, 2002).

Also, the number of women in Parliament takes into account differences in the number of women in decision-making jobs. In mixed organizations, staff positions are affected by gender. Women belong to the vulnerable category and are mainly represented in the informal sector than in the formal sector. According to the exclusion of social protection is a phenomenon strongly related to gender. Women who manage micro-enterprises are confronted with difficulties relating to gender because of social and cultural influences.

#### 3.2.2. Nature and Sources of Data

The World Bank on its website<sup>9</sup> publishes the macroeconomic indicators of each country. All our data comes from this website. We chose to use data of the year 2012 because they are relatively more available for many countries. Despite the enlightenment that the different approaches to the measurement of the informal economy bring, the data of these studies are different.

## 4. RESULTS AND INTERPRETATION: INDICES OF THE MEASUREMENT OF INFORMALITY

### 4.1. The Elementary Index of Informality

Figures<sup>10</sup> 1 and 2 see Appendices 1 and 2 show the results of a principal components analysis<sup>11</sup> on 189 countries. Factorial axis 1 opposes financial variables (domestic credit granted by the financial sector in % of the GDP; number of deposits in banks for 1000 adults;...), welfare variables (GDP per capita; consumption of electricity per capita; final household consumption...) against variables measuring administrative and tax difficulties (total number of taxes as a % of commercial profits; time in hours to prepare and pay taxes; the cost to begin an activity). The other variables that characterize the first factor are the dependency ratio of the elderly; the rate of failure in primary and secondary education. The value-added in agriculture per capita is strongly correlated with the welfare indicators.

The factorial axis 2 is primarily made of variables which characterise unemployment. Certain variables were used for illustrative purposes since their contribution to the independent factors is very low (gross fixed capital formation; time in days to register a good; time in days to enforce a contract; time in days to begin a commercial

<sup>9</sup><http://data.worldbank.org/data-catalog/world-development-indicator>.

<sup>10</sup> For descriptive analysis, we use data analysis techniques. Data analysis is a set of techniques that enable the extraction of statistical information contained in large tables. These are geometrical methods method that give an 'x-ray' of data. They describe the data globally and enable the obtaining of its internal structure in terms of the different axes or homogenous trends. These are multi-dimensional descriptive methods that enable the management of a large quantity of data and variables.

<sup>11</sup> The principal components analysis is done using version 5.5 of the SPAD software.

activity; and the proportion of the women having a seat in parliament). Other variables are also used as for illustrative purposes because they are strongly correlated with other variables measuring the welfare of the populations (final government consumption; gross national income per capita, and healthcare expenditure per capita). These variables are redundant. They contain information which is related to variables measuring welfare (GDP per capita; ...).

Figure 2 in Appendix 2 shows the distribution of the countries following the main factorial axes 1 and 2. The first group of countries<sup>12</sup> (Chad; South-Sudan; Democratic Republic of Congo; Liberia; Zimbabwe; Guinea-Bissau; Niger; Togo; Mali; Mozambique; Sierra Leone; Madagascar; Ethiopia; Rwanda; Burundi; ...) is made up of very poor countries with very low gross income per capita. These are mostly countries of Sub-Saharan Africa where the cost of regulation and administrative procedures are highest. They are also countries where the rates of failure to in primary and secondary education are highest and where the rate of adult dependency is highest. In these countries, the indicators of welfare are low (GDP per capita, consumption of electricity per capital ;...).

The second group of countries (Namibia; Tajikistan; Algeria; Philippines; Swaziland; Belize; El Salvador; Bangladesh; Nicaragua; Nigeria; Pakistan; Paraguay; India; Ghana; Kosovo) is made up of relatively poor countries where the regulatory burden is particularly high (time to begin a commercial activity and time to prepare and pay taxes).

The third group of countries (Poland; Jordan; Lithuania; Turkey; Hungary; Romania; Maldives; Lithuania; Seychelles; Slovakia; Albania; the Caribbean islands; Georgia; Greece; Armenia; ...) is made up of countries with intermediate incomes where the global unemployment rate and youth unemployment are particularly high. In this group of countries, the regulatory burden and tax authorities are relatively low.

The fourth group of countries (Norway; Luxembourg; Japan; Australia; Hong-Kong; Qatar; Denmark; The United States; Sweden; Canada; Austria; Germany; New Zealand; Netherlands; Belgium; Finland;...) is made up of developed and OECD countries. They are countries where the average gross income per capita was about 36 000 US \$ during the year 2012. In this group of countries, the population easily has access to electricity, banking services and domestic credit. In addition to these characteristics, it should be noted that in these countries, the regulatory burden and administrative procedures are low; the unemployment rate is also relatively low.

Figure 2 see Appendix 2 classifies countries according to their level of informality; going from the left to the right, we move from countries with a highly informal economy towards countries that are less informal. The grouping of the countries is done according to the criteria which characterize informality.

Each variable of Figure 1 in the appendix describes a dimension of informality. Thus, informality arises in several ways, according to whether one is in the poor countries of sub-Saharan Africa, in countries with intermediate incomes or the developed countries of the OECD. According to the exploratory analysis, the traces of informality are more present in the poor countries of sub-Saharan Africa than in the developed and OECD countries.

In addition, we build for each variable characterizing a dimension of informality, an elementary indicator of informality. The elementary indicators are then too assembled to build an aggregate index of informality.

The Economy of a country is a priori informal for the criterion **X** if its elementary index of informality for this criterion is higher than the median. The economy of a country is a priori informal<sup>13</sup> if it is informal for more than half of the selected criteria of informality. We transform each variable into an elementary index of informality.

<sup>12</sup> We chose to partition the countries in four groups shown by the yellow marks in Figure 2.

<sup>13</sup> The definition and measurement informality depends on the chosen criteria.

When a variable  $X$  is such that its increase is believed to increase the risk of informality, then the elementary index

of informality of the country  $i$ , for the criterion  $X$  is given by<sup>14</sup>: 
$$\frac{X_i}{\max_{j=1,2,\dots,N}(X_j)}$$

When a variable is such that its increase is likely to reduce the risk of informality, then the elementary index of

informality of the country for the criterion is equal to: 
$$1 - \frac{X_i}{\max_{j=1,2,\dots,N}(X_j)}$$

Table 1 gives a list of variables that describe informality. In this paragraph, we only retain the variables that are most significant in the exploratory analysis. These are the variables<sup>15</sup> best represented on the main factorial axes or which contribute the greatest shares to the formation of the factors. Among the variables likely to increase informality, we can cite: total taxes as a percentage of commercial profits; the cost to start a commercial activity; the rate of failure in primary education; the adult dependency ratio; and the youth unemployment rate. In fact, these are globally the variables of Figure 1 that are directed to the left or upwards.

Among the variables likely to decrease the level of informality, we can cite: GDP per capita; the consumption of electricity per capita; the domestic credit granted by banks as a percentage of GDP; the value added in agriculture per capita; and per capita money supply. These are generally the variables in Figure 1 that are directed towards the left. Thus, the basic index of informality for a country and a criterion is always between 0 and 1. The basic indices are built so that their increase increases the risk of informality. When a basic index of informality is close to 0, this means that the traces of informality for this criterion are weak; when it is close to 1, then the traces of informality are high for this criterion.

We say that the economy of a country is a priori informal for the criterion  $X$  if its basic index of informality for this criterion is higher than the median. The economy of a country is a priori informal if it is informal for more than of the selected criteria of informality<sup>16</sup>. Table 2 gives an alphabetical list of the most informal countries a priori. The score 1 for a criterion means that the country is a priori informal for this criterion; 0 mean that the country is not a priori informal for the indicated criterion. For the 10 criteria selected, the countries that carry more than 5 times the score 1 are a priori informal.

Table 2 shows that in the a priori most informal countries, the youth unemployment rate is generally low. Generally, these countries are a priori informal for all the criteria, except the youth unemployment rate. The least a priori informal countries for the criteria selected are: Brazil, Canada, Chile, Denmark, Dominican Islands, Hong-Kong, Iceland, Korean Republic, Luxembourg, Malaysia, Netherlands, Qatar, Singapore, Switzerland, Turkey, United Arab Emirates, and the United States. These are mainly countries located in the South-east of Figure 2.

---

<sup>14</sup> Max refers to the maximum observed for the variable  $X; j = 1, 2, \dots, N$ , denote the countries;  $N$  is the number of countries.

<sup>15</sup> These variables are listed in Table 2

<sup>16</sup> The definition and quantification of informality depends on the criteria chosen.

Tablea-2. The a priori most informal countries.

Country Name	GDP per capita	Electric power (kWh per capita)	Agriculture value added per worker	Age dependency ratio	Money (M2) per capita	Cost of business start-up procedures	Domestic credit provided by financial sector	Repeats primary school	Total tax rate (% of commercial profits)	Unemployment, youth total	Number of informality scores	A Priori informality
Benin	1	1	1	1	1	1	1	1	1	0	9	1
Burkina Faso	1	1	1	1	1	1	1	1	1	0	9	1
Cameroon	1	1	1	1	1	1	1	1	1	0	9	1
RCA	1	1	1	1	1	1	1	1	1	0	9	1
Chad	1	1	1	1	1	1	1	1	1	0	9	1
Comoros	1	1	1	1	1	1	1	1	1	0	9	1
Democratic Republic of Congo	1	1	1	1	1	1	1	1	1	0	9	1
Cote d'Ivoire	1	1	1	1	1	1	1	1	1	0	9	1
Guinea	1	1	1	1	1	1	1	1	1	0	9	1
Guinea - Bissau	1	1	1	1	1	1	1	1	1	0	9	1
Haiti	1	1	1	1	1	1	1	1	1	0	9	1
Kenya	1	1	1	1	1	1	1	1	1	0	9	1
Mali	1	1	1	1	1	1	1	1	1	0	9	1
Mauritania	1	1	1	1	1	1	1	0	1	1	9	1
Small states	0	1	1	1	1	1	1	1	1	1	9	1
Togo	1	1	1	1	1	1	1	1	1	0	9	1
Republic of Yemen	1	1	1	1	1	1	1	1	0	1	9	1

#### 4.2. The Aggregate Index of Informality

The logistic regression is done with version 8.0 of the “Eviews” econometrics software. The variable to be explained is informality. The explanatory variables are: GDP per capita; the consumption of electricity in kWh per capita; the value added in agriculture; the rate of dependence of adults; money and quasi money per capita; the cost of procedures to begin an economic activity; domestic credit granted by the financial sector as a percentage of the GDP; the rate of failure in primary education; the global rate of taxation as a percentage of commercial profits; and the youth unemployment rate.

Table 3, gives the results of the logistic regression on the set of 189 countries selected. Note that the coefficients of this regression can be interpreted as the weights of the variables which are associated to them. Knowing the coefficients of the model, it becomes easy to envisage the probability (which we call also degree) of informality of a country given its characteristics. If a country does not form part of our sample, we can, knowing its characteristics (GDP per capita, consumption of electricity per capital, ...), calculate its degree of informality. The results of this regression are presented in Table 3.

Table-3, Results of the logistic regression using Eviews 8.0.

Dependent Variable: INFORMILITE_A_PRIORI				
Method: ML - Binary Logit (Quadratic hill climbing)				
Sample: 1 189				
Included observations: 189				
Convergence achieved after 8 iterations				
Covariance matrix computed using second derivatives				
Variable	Coefficient	Standard. Error	z-Statistic	Prob.
Gdp_Per_Capita	1.461805	0.432742	3.378005	0.0007
Electric_Power_per_Capit	0.028358	0.035549	0.797707	0.4250
Money_M2_Per_Capita	2.236586	0.552650	4.047018	0.0001
Agricult_Value_Added_per	1.779468	0.482526	3.687818	0.0002
Domestic_Credit_Financi	0.279028	0.149033	1.872251	0.0612
Repeaters_Primary_School	0.372993	0.133375	2.796562	0.0052
Total_Tax_Rate	0.986562	0.195872	5.036757	0.0000
Cost_Business_Start_Up_P	1.533072	0.295666	5.185154	0.0000
Unemployment_of_Youth	0.219015	0.109656	1.997292	0.0458
Age_Dependency_Ratio	1.070668	0.285380	3.751731	0.0002
C	6.523453	0.926498	7.040982	0.0000
McFadden R-squared	0.889027	Mean dependent variable		0.408377
S.D. dependent var	0.492825	S.E. of regression		0.153073
Akaike info criterion	0.265278	Sum squared resid		4.217665
Schwarz criterion	0.452581	Log likelihood		-14.33401
Hannan-Quinn criter.	0.341144	Deviance		28.66803
Restr. Deviance	258.3322	Restr. log likelihood		-129.1661
LR statistic	229.6642	Avg. log likelihood		-0.075047
Prob(LR statistic)	0.000000			
ObswithDep=0	113	Total observations		189
ObswithDep=1	76			

All the coefficients of the model are positive, which reflects coherence in the construction of the model (the elementary indices of informality are built so that their increase leads to an increase in the level of informality). The constant  $\hat{c} = 6,523$  reflects the threshold of informality. The economy in a country is declared a posteriori informal (from the results of the model) if its score of informality is higher than  $\hat{c}$ . The model is globally acceptable



see Table 3, statistics of the likelihood ratio) although some parameters<sup>17</sup> are not significantly different from 0. Table 4 below shows the different levels of informality of each country.

Table-4. The degree of informality of some countries<sup>18</sup>.

Name of the Country	Number of scores of informality	Probability of informality in %	Score of informality	dimension of risk	Informality a priori
Afghanistan	7	70.2	7,38	2.36	1
Albania	2	42.0	6,20	0.72	0
Algeria	6	55.2	6,73	1.23	1
Angola	6	77.8	7,78	3.50	1
Antigua	2	21.6	5,23	0.27	0
Argentina	3	43.2	6,25	0.76	0
Armenia	4	40.2	6,13	0.67	0
Australia	1	0.4	1,06	0.00	0
Austria	1	5.4	3,66	0.06	0
Azerbaijan	2	40.5	6,14	0.68	0
The Bahamas	2	3.4	3,17	0.03	0
Baharin	1	6.9	3,92	0.07	0
Bangladesh	8	66.9	7,23	2.03	1
Barbados	2	28.4	5,60	0.40	0
Belarus	2	37.6	6,02	0.60	0
Belgium	1	5.2	3,62	0.05	0
Belize	6	49.9	6,52	0.99	1
Benin	9	80.6	7,95	4.16	1

Table 2 is primarily descriptive. It gives the list of the countries with marks of informality. However, it does not inform on the extent of the informality and does not take into account the weight of each criterion of informality. Table 4 combines the elementary indices of informality and the associated weights to calculate the scores, the dimensions and the probabilities (degrees) of informality.

The score, dimension of informality, and the probability of informality are equivalent criteria making it possible to classify the economies according to their degree of informality. In Table 4, we note that Bangladesh which has a priori 8 marks of informality on 10 is more informal a priori than Angola that has 6 marks of informality a priori. When we take into account of size of informality and their weights, we find that Angola is a posteriori (following the results of the logistic model) more informal than Bangladesh.

According to the results of the logistic model, the 10 most informal countries in decreasing order are: the Democratic republic of Congo, Chad, Gambia, Haiti, Comoros, Central African Republic, South-Sudan, Guinea, Malawi, and Togo. In the same manner, the 10 least informal countries a posteriori are (by order ascending of informality): Singapore, Qatar, Denmark, United States, Japan, Australia, Macao, Iceland, Sweden, and Switzerland.

<sup>17</sup> The coefficient of electricity consumption per head is not statistically different from 0; the criterion: domestic credit from banks is not statistically significant at the 5% level.

<sup>18</sup> The complete list is in the table in Appendix 3.

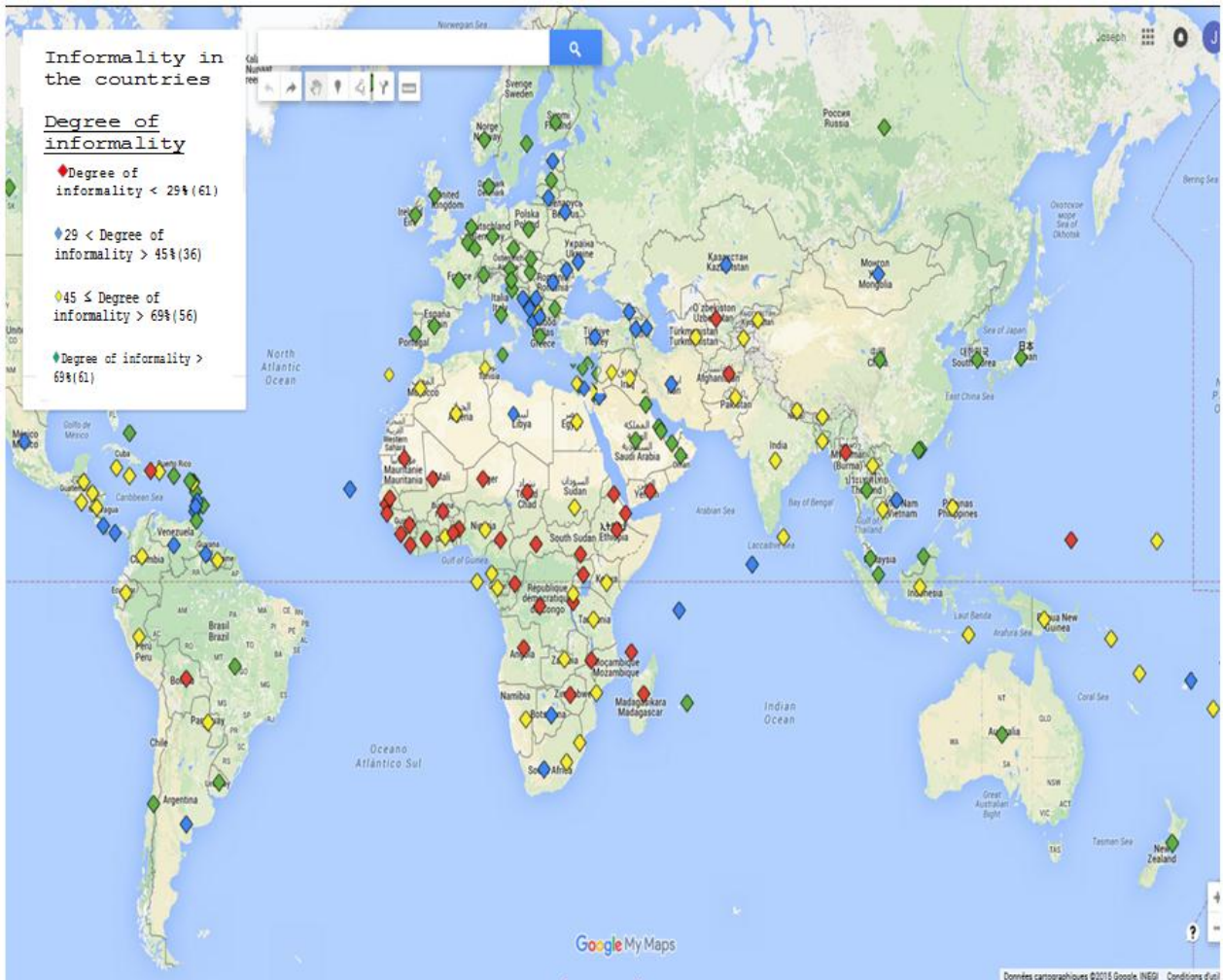


Figure-1. Graphical representation of countries according to their levels of informality.

Figure 1 describes the countries according to their degree of informality. The red marks show highly informal countries (level of informality is higher than 69%). They are mainly the countries of Sub-Saharan Africa and some countries of South Asia and Latin America. The yellow marks show countries where the degree of informality is relatively high (degree of informality ranging between 45 and 69%). These are located mainly in Africa, Latin America and Asia. Blue marks denote countries where the degree of informality is relatively low (ranging between 29 and 45%). It should be noted that these countries are dispersed on the planet, with a small majority in Europe. The green marks denote countries that are slightly informal (degree of informality lower than 29%). They are mainly the countries of Europe and North America.

Table 5 below gives the elasticities of the risk of informality relative to the factors of informality for some countries with high income. In this table, we notice that the elasticity of the degree of informality relative to the consumption of electricity is very low. If the consumption of electricity decreases, the degree of informality does not increase significantly. On the other hand, if the basic index of informality related to the money supply per capita increases by 1% (i.e. the quantity of money per capita decreases), then the degree of informality will increase by approximately 2% in the Netherlands, Finland, Belgium, Germany, Canada, Austria and France. Elasticities indicate the factors likely to make informality very quickly.

Table-5. Elasticity of informality relative to the determinants of informality.

Country Names	GDP per Capita	Electric power per Capita	Agriculture value added worker	Age dependency ratio	money (M2) capita	Cost of business start-up	Domestic credit y financial sector)	Repeaters primary school	Total tax rate	Unemployment, youth
U. Kingdom	0,26	0,01	0,49	0,50	0,00	0,00	0,01	0,00	0,34	0,09
Hong Kong	0,39	0,01	0,74	0,32	0,00	0,01	0,00	0,01	0,22	0,04
Korea, Rep.	0,74	0,00	0,69	0,35	0,00	0,08	0,06	0,00	0,33	0,04
Emirates	0,24	0,00	1,24	0,16	0,28	0,03	0,17	0,00	0,14	0,04
Brunei	0,24	0,00	1,10	0,37	0,18	0,06	0,26	0,00	0,15	0,05
Luxembourg	0,00	0,00	0,05	0,42	1,91	0,01	0,00	0,00	0,19	0,08
Norway	0,00	0,00	0,00	0,48	1,90	0,01	0,11	0,01	0,39	0,04
Lebanon	1,14	0,02	0,19	0,44	0,48	0,18	0,03	0,10	0,29	0,08
Israel	0,50	0,01	1,05	0,57	0,32	0,02	0,16	0,02	0,27	0,05
Netherlands	0,02	0,01	0,00	0,47	2,04	0,03	0,00	0,02	0,37	0,04
Bahamas,	0,79	0,01	0,18	0,38	0,95	0,05	0,13	0,00	0,43	0,13
Finland	0,07	0,00	0,00	0,49	2,03	0,01	0,05	0,00	0,39	0,08
Canada	0,00	0,00	0,68	0,41	1,99	0,00	0,08	0,02	0,19	0,06
Belgium	0,14	0,01	0,00	0,47	2,00	0,03	0,12	0,03	0,54	0,08
Germany	0,17	0,01	0,24	0,47	2,01	0,02	0,06	0,01	0,43	0,03
Austria	0,05	0,00	0,32	0,44	1,99	0,02	0,08	0,03	0,49	0,04
S. Arabia	0,69	0,00	0,66	0,43	1,19	0,03	0,26	0,02	0,14	0,12
France	0,25	0,01	0,00	0,51	1,99	0,00	0,07	0,02	0,62	0,10

## 5. CONCLUSION

This study has as objective to characterize informal economies. It identifies the factors that determine informality and of synthesizes them into indices measuring informality. We use data on 189 countries and analyze only the data for the year 2012. We used two approaches to examine informality in the economies. The first approach is exploratory and enables us to classify the countries into four groups according to the criteria which best describes informality in each group. The second approach is based on the results of the first and uses the techniques of logistic regression to build a synthetic indicator which measures the degree of informality of the economy of each country. The results of the two approaches show that the marks of informality are more visible in developing countries of Sub-Saharan Africa than in developed countries of the OECD. Moreover, the degree of informality in the countries is related to their level of development. Thus, developing countries are generally those where the degree of informality is highest while developed countries generally have lower degrees of informality. This study made it possible to organize the countries in groups according to the variables which characterize informality and to visualize on maps the countries according to their degree of informality. The measurement of the size of the informal economy is of a major importance because it makes it possible to correct the macroeconomic indicators such as the GDP per capita, the cost of living, poverty indices, and the index of human development. The taking into account of the informal economy also makes it possible to better target economic diagnosis and decisions to evaluate the impact of social policies and tax decisions.

**Funding:** This study received no specific financial support.

**Competing Interests:** The authors declare that they have no competing interests.

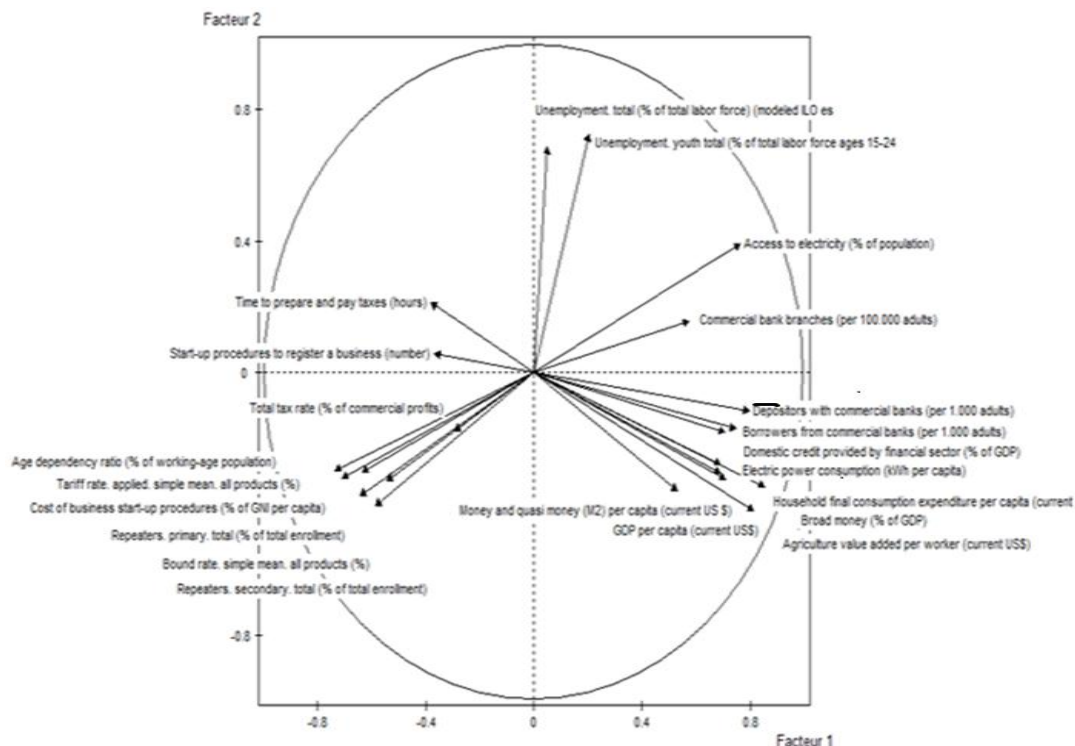
**Acknowledgement:** All authors contributed equally to the conception and design of the study.

## REFERENCES

- Bacchetta, M., Ernst, E., & Bustamante, J. P. (2009). *Globalization and informal jobs in developing countries*. Geneva: International Labour Organization.
- Cagan, P. (1958). The demand for currency relative to the total money supply. *Journal of Political Economy*, 66(4), 303-328. Available at: <https://doi.org/10.1086/258056>.
- Contini, B. (1981). Labour market segmentation and the development of the parallel economy: The Italian experience. *Oxford Economic Papers*, 33(3), 401-412. Available at: <https://doi.org/10.1093/oxfordjournals.oep.a041515>.
- Davidson, R., & MacKinnon, J. G. (1993). *Estimation and inference in econometrics*. New York: Oxford University Press.
- De Soto, H. (1989). *The other path*. New York: Harper and Row.
- Feige, E. L. (1979). How big is the irregular economy? *Challenge*, 22(5), 5-13. Available at: <https://doi.org/10.1080/05775132.1979.11470559>.
- Frey, B. S., & Weck, H. (1983). Estimating the shadow economy: A naive approach. *Oxford Economic Papers*, 35(1), 23-44. Available at: <https://doi.org/10.1093/oxfordjournals.oep.a041582>.
- Gutmann, P. M. (1977). The subterranean economy. *Financial Analysts Journal*, 33(6), 26-27.
- Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development: A two-sector analysis. *The American Economic Review*, 60(1), 126-142.
- Helberger, C., & Knepel, H. (1988). How big is the shadow economy?: A re-analysis of the unobserved-variable approach of BS Frey and H. Weck-Hannemann. *European Economic Review*, 32(4), 965-976.
- ILO. (2002). *Decent work and the informal economy: Agenda item 6*. Paper presented at the Geneva, International Labour Conference, 90th Session.
- Kaufmann, D., & Kaliberda, A. (1996). Integrating the unofficial economy into the dynamics of post-socialist economies: A framework of analysis and evidence, in Kaminski and Barlomiej, *Economic Transition in Russia and the new states of Eurasia*. *International Politics of Eurasia*, 8, 81-120.
- Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The Manchester School*, 22(2), 139-191.

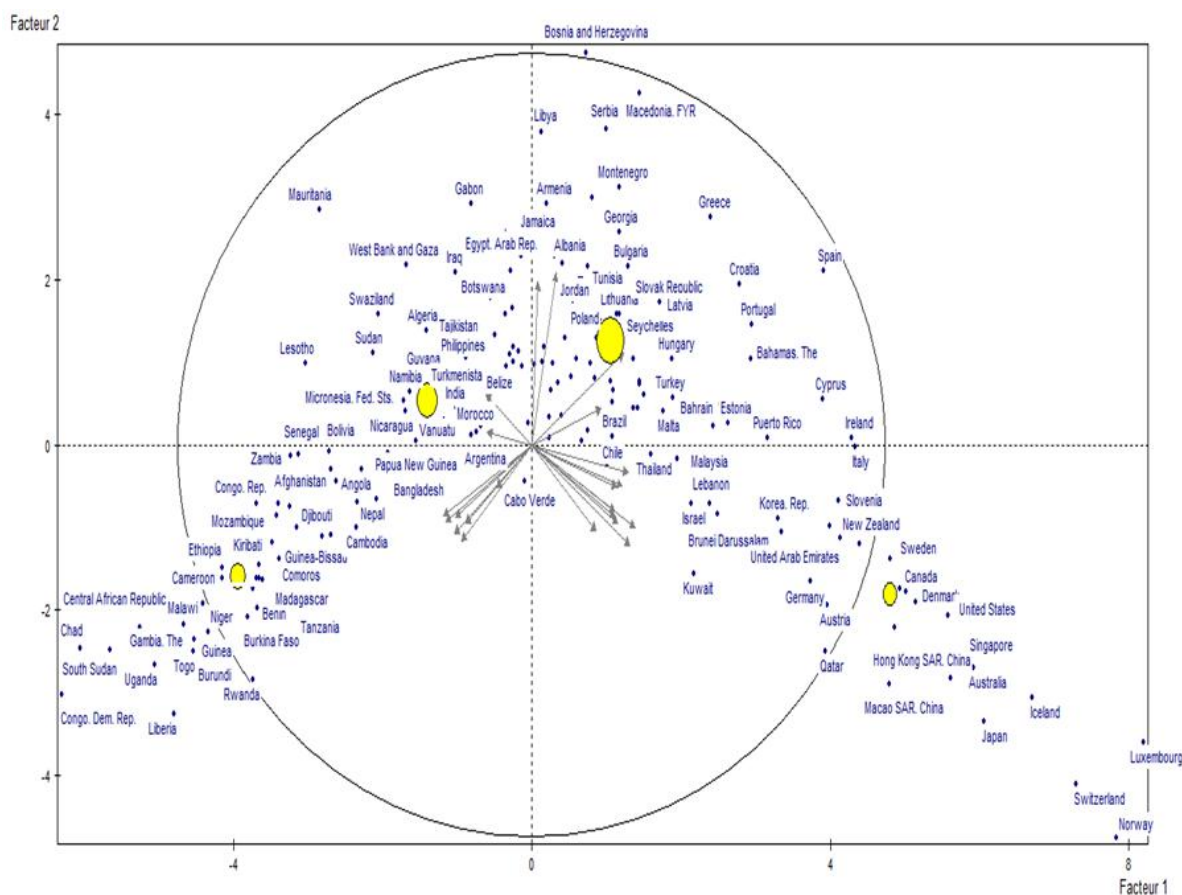
- Maldonado, C. (2001). *The informal economy in French-speaking Africa: Structure, dynamics and politics*. Geneva: B.I.T.
- Maloney, W. F. (2004). Informality revisited. *World Development*, 32(7), 1159-1178. Available at: <https://doi.org/10.1016/j.worlddev.2004.01.008>.
- Maurizio, R. (2012). *Labour informality in Latin America: The case of Argentina, Chile, Brazil and Argentina, BWPI Working Papers 165/2012*. Manchester, UK: Brooks World Poverty Institute, University of Manchester.
- Muheme, G. B. (1995). *The informal economy in the face of the evolution of dualism. Theoretical conceptualization and illustration by studying the agro-economic route of the Mboga sector*. Quebec, Thesis, Laval University.
- Portes, A., Castells, M., & Benton, L. A. (1989). *The informal economy: Studies in advanced and less developed countries*. Baltimore: The John Hopkins University Press.
- Schneider, F. (2004). The size of the shadow economies of 145 countries all over the world: First results over the period 1999 to 2003. IZA Discussion Paper 1431.
- Schneider, F. (2000). Dimensions of the shadow economy. *The Independent Review*, 5(1), 81-91.
- Tanzi, V. (1980). The underground economy in the United States: Estimates and implications. *National Labor Bank*, 135(4), 427-453.
- Tanzi, V. (1982). *The underground economy in the United States and abroad*: Free Press.
- Thomas, A. (2002a). Econometrics of qualitative variables (pp. 56-57): Dunod.
- Thomas, A. (2002b). Econometrics of qualitative variables (pp. 60-61): Dunod.
- Walther, R. (2006). Vocational training in the informal sector: Report on the field survey in Cameroon. Working Document of the French Development Agency, 17.

## APPENDICES



Appendix-1. Representation of the variables on the 1st and 2<sup>nd</sup> factorial axes Figure 1.





Appendix-2. Description of countries and variables on the 1st and 2<sup>nd</sup> factorial axes Figure 2.

Appendix-3. Determination of the level of informality of countries.

Country Name	Sum of indicators of informality	A priori Informality	SCORE	ODD	Probability of informality
Afghanistan	7	1	7,38	2,36	0,70
Albania	2	0	6,20	0,72	0,42
Algeria	6	1	6,73	1,23	0,55
Angola	6	1	7,78	3,50	0,78
Antigua and Barbuda	2	0	5,23	0,27	0,22
Argentina	3	0	6,25	0,76	0,43
Armenia	4	0	6,13	0,67	0,40
Australia	1	0	1,06	0,00	0,00
Austria	1	0	3,66	0,06	0,05
Azerbaijan	2	0	6,14	0,68	0,40
Bahamas, The	2	0	3,17	0,03	0,03
Bahrain	1	0	3,92	0,07	0,07
Bangladesh	8	1	7,23	2,03	0,67
Barbados	2	0	5,60	0,40	0,28
Belarus	2	0	6,02	0,60	0,38
Belgium	1	0	3,62	0,05	0,05
Belize	6	1	6,52	0,99	0,50
Benin	9	1	7,95	4,16	0,81
Bhutan	4	0	6,48	0,96	0,49
Bolivia	7	1	7,41	2,43	0,71
Bosnia and Herzegovina	3	0	6,25	0,76	0,43
Botswana	4	0	6,30	0,80	0,44
Brazil	0	0	5,17	0,26	0,21
Brunei Darussalam	1	0	2,44	0,02	0,02
Bulgaria	1	0	5,25	0,28	0,22
Burkina Faso	9	1	7,42	2,44	0,71

Burundi	8	1	7,55	2,80	0,74
Cabo Verde	4	0	6,30	0,80	0,45
Cambodia	7	1	7,03	1,66	0,62
Cameroon	9	1	7,34	2,27	0,69
Canada	0	0	3,61	0,05	0,05
Caribbean small states	2	0	6,37	0,86	0,46
Central African Republic	9	1	8,36	6,31	0,86
Chad	9	1	8,87	10,50	0,91
Chile	0	0	4,86	0,19	0,16
China	1	0	5,13	0,25	0,20
Colombia	3	0	6,40	0,88	0,47
Comoros	9	1	8,38	6,38	0,86
Congo, Dem. Rep.	9	1	9,40	17,68	0,95
Congo, Rep.	9	1	7,66	3,13	0,76
Costa Rica	1	0	6,03	0,61	0,38
Cote d'Ivoire	9	1	7,88	3,87	0,79
Croatia	1	0	4,43	0,12	0,11
Cyprus	2	0	4,75	0,17	0,15
Czech Republic	1	0	4,74	0,17	0,14
Denmark	0	0	0,87	0,00	0,00
Djibouti	8	1	7,73	3,33	0,77
Dominica	0	0	5,62	0,41	0,29
Dominican Republic	6	1	6,39	0,88	0,47
Ecuador	5	0	6,50	0,98	0,49
Egypt, Arab Republic	5	0	6,57	1,05	0,51
El Salvador	6	1	6,68	1,17	0,54
Equatorial Guinea	6	1	6,74	1,25	0,55
Eritrea	8	1	7,67	3,14	0,76
Estonia	3	0	5,79	0,48	0,32
Ethiopia	8	1	7,75	3,41	0,77
Fiji	2	0	6,23	0,74	0,43
Finland	1	0	3,23	0,04	0,04
France	4	0	3,82	0,07	0,06
Gabon	6	1	6,56	1,04	0,51
Gambia, The	8	1	8,51	7,26	0,88
Georgia	4	0	6,21	0,73	0,42
Germany	1	0	3,63	0,06	0,05
Ghana	6	1	6,86	1,41	0,58
Greece	3	0	5,52	0,37	0,27
Grenada	1	0	5,89	0,53	0,35
Guatemala	8	1	6,98	1,58	0,61
Guinea	9	1	8,16	5,16	0,84
Guinea-Bissau	9	1	7,41	2,42	0,71
Guyana	4	0	6,29	0,79	0,44
Haiti	9	1	8,44	6,77	0,87
Honduras	5	0	6,74	1,24	0,55
Hong Kong SAR, China	0	0	1,76	0,01	0,01
Hungary	2	0	5,44	0,34	0,25
Iceland	0	0	1,08	0,00	0,00
India	6	1	6,75	1,25	0,56
Indonesia	7	1	6,65	1,14	0,53
Iran, Islamic Republic.	3	0	6,07	0,63	0,39
Iraq	5	0	6,75	1,26	0,56
Ireland	1	0	4,52	0,14	0,12
Israel	1	0	3,06	0,03	0,03
Italy	3	0	4,06	0,09	0,08
Jamaica	3	0	6,37	0,86	0,46
Japan	1	0	0,93	0,00	0,00
Jordan	3	0	6,04	0,61	0,38
Kazakhstan	1	0	5,65	0,42	0,29



Kenya	9	1	7,23	2,03	0,67
Kiribati	7	1	7,20	1,97	0,66
Korea, Rep.	0	0	2,32	0,01	0,01
Kosovo	5	0	6,46	0,94	0,48
Kuwait	1	0	1,59	0,01	0,01
Kyrgyz Republic	4	0	6,57	1,04	0,51
Lao PDR	6	1	6,74	1,24	0,55
Latvia	1	0	5,32	0,30	0,23
Lebanon	2	0	3,05	0,03	0,03
Lesotho	8	1	6,93	1,49	0,60
Liberia	8	1	7,45	2,52	0,72
Libya	3	0	5,80	0,48	0,33
Lithuania	3	0	5,75	0,46	0,32
Luxembourg	0	0	2,72	0,02	0,02
Macao SAR, China	1	0	1,08	0,00	0,00
Macedonia, FYR	2	0	5,64	0,41	0,29
Madagascar	8	1	7,46	2,56	0,72
Malawi	8	1	8,01	4,43	0,82
Malaysia	0	0	4,66	0,16	0,13
Maldives	1	0	5,95	0,56	0,36
Mali	9	1	7,90	3,95	0,80
Malta	2	0	5,27	0,28	0,22
Marshall Islands	6	1	7,06	1,70	0,63
Mauritania	9	1	7,54	2,76	0,73
Mauritius	1	0	5,14	0,25	0,20
Mexico	1	0	5,86	0,51	0,34
Micronesia, Fed..	8	1	7,60	2,92	0,74
Moldova	4	0	6,27	0,78	0,44
Mongolia	2	0	6,06	0,63	0,39
Montenegro	1	0	5,85	0,51	0,34
Morocco	4	0	6,35	0,84	0,46
Mozambique	7	1	7,22	2,01	0,67
Myanmar	7	1	7,47	2,59	0,72
Namibia	4	0	6,50	0,98	0,49
Nepal	7	1	6,93	1,50	0,60
Netherlands	0	0	3,10	0,03	0,03
New Zealand	1	0	3,95	0,08	0,07
Nicaragua	7	1	7,28	2,13	0,68
Niger	8	1	7,93	4,06	0,80
Nigeria	7	1	7,32	2,22	0,69
Norway	1	0	3,02	0,03	0,03
Oman	2	0	4,92	0,20	0,17
Pakistan	7	1	6,83	1,36	0,58
Palau	3	0	6,42	0,90	0,47
Panama	1	0	5,69	0,43	0,30
Papua New Guinea	7	1	6,76	1,27	0,56
Paraguay	6	1	6,62	1,10	0,52
Peru	3	0	6,33	0,82	0,45
Philippines	5	0	6,68	1,17	0,54
Poland	2	0	5,60	0,40	0,28
Portugal	3	0	5,58	0,39	0,28
Puerto Rico	3	0	4,10	0,09	0,08
Qatar	0	0	0,85	0,00	0,00
Romania	2	0	5,82	0,50	0,33
Russian Federation	1	0	5,32	0,30	0,23
Rwanda	6	1	7,02	1,65	0,62
Samoa	3	0	6,39	0,87	0,47
Sao Tome and Principe	7	1	7,16	1,89	0,65
Saudi Arabia	2	0	3,75	0,06	0,06
Senegal	8	1	7,40	2,41	0,71

Serbia	1	0	6,22	0,74	0,42
Seychelles	2	0	5,66	0,42	0,30
Sierra Leone	8	1	7,45	2,54	0,72
Singapore	0	0	0,84	0,00	0,00
Slovak Republic	3	0	5,53	0,37	0,27
Slovenia	2	0	4,01	0,08	0,07
Small states	9	1	6,98	1,58	0,61
Solomon Islands	7	1	6,98	1,57	0,61
South Africa	1	0	5,67	0,43	0,30
South Asia	5	0	6,71	1,21	0,55
South Sudan	8	1	8,27	5,71	0,85
Spain	2	0	4,20	0,10	0,09
Sri Lanka	6	1	6,75	1,25	0,56
St. Kitts and Nevis	3	0	4,99	0,22	0,18
St. Lucia	1	0	5,86	0,51	0,34
St. Vincent and the Grenadines	1	0	6,13	0,68	0,40
Sudan	7	1	7,03	1,67	0,62
Suriname	3	0	6,38	0,86	0,46
Swaziland	8	1	7,08	1,75	0,64
Sweden	2	0	1,29	0,01	0,01
Switzerland	0	0	1,40	0,01	0,01
Tajikistan	6	1	7,32	2,21	0,69
Tanzania	8	1	7,28	2,12	0,68
Thailand	1	0	5,59	0,39	0,28
Timor-Leste	7	1	6,83	1,36	0,58
Togo	9	1	8,01	4,40	0,81
Tonga	4	0	6,46	0,93	0,48
Trinidad and Tobago	2	0	5,08	0,24	0,19
Tunisia	4	0	6,34	0,83	0,45
Turkey	0	0	5,67	0,43	0,30
Turkmenistan	4	0	6,48	0,96	0,49
Uganda	8	1	7,82	3,67	0,79
Ukraine	2	0	6,23	0,74	0,43
United Arab Emirates	0	0	2,34	0,02	0,01
United Kingdom	1	0	1,72	0,01	0,01
United States	0	0	0,91	0,00	0,00
Uruguay	3	0	5,56	0,38	0,28
Uzbekistan	6	1	7,43	2,49	0,71
Vanuatu	4	0	6,40	0,88	0,47
Venezuela, RB	2	0	5,90	0,54	0,35
Vietnam	3	0	6,25	0,76	0,43
West Bank and Gaza	7	1	7,14	1,85	0,65
Yemen, Rep.	9	1	7,33	2,23	0,69
Zambia	7	1	7,07	1,73	0,63
Zimbabwe	7	1	7,54	2,77	0,73

*Views and opinions expressed in this article are the views and opinions of the author(s), Asian Journal of Economic Modelling shall not be responsible or answerable for any loss, damage or liability etc. caused in relation to/arising out of the use of the content.*