

Modeling Factors Influencing Employment in Informal Sector Participation in Rwanda (2017) Using Logistic Regression Approach

Pacifique UWAMAHORO¹, Joseph K. Mung'atu²

¹ Jomo Kenyatta University of Agriculture and Technology, Kenya

Abstract: *The researcher observed that the employment in informal sector represents roughly third quarter in many developing countries, from that the researcher anticipated that once the informal sector is taken advantage of, it might reduce the unemployment rate, which is being a major issue to be addressed. This research aimed at modelling different factors influencing employment in informal sector participation in Rwanda. it was undertaken using secondary data "Rwanda Labour force survey August 2017 data" which are the data collected in August 2017 by NISR in a survey called Labour Force Survey.*

Keywords: Informal sector, Labor force, employment, unemployment

1. Introduction

The ILO launched the concept of the informal sector three decades ago. Since then, it has done more work on both the concept and the underlying social problem than any other single institution. The informal sector can be described as a grey market in Labour.

Other concepts which can be characterized as informal sector can include the black market (shadow economy, underground economy), agorism and system D. Associated idioms include "under the table", "off the books" and "working for cash". (Internation Labour Organisation, n.d.)

Although the informal sector makes up a significant portion of the economies in developing countries it is often stigmatized as troublesome and unmanageable. However, the informal sector provides critical economic opportunities for the poor, and has been expanding rapidly since the 1960s. As such, integrating the informal economy into the formal sector is an important policy challenge.

Persons employed in the informal sector can be identified as follows:

An employee or contributing family worker is employed in the informal sector if:

the number of persons working in the enterprise where that person works is less than some specified cut-off (e.g. five); and either the enterprise is not formally registered (for purposes of tax, social security or some form of trade licence) or it operates in certain types of informal places such as at or near the home, in the street or in some moveable location.

An employer, own-account worker or member of a producers' cooperative is employed in the informal sector if the enterprise they own is not formally registered (for purposes of tax, social security or some form of trade licence) or does not have a bank account separate from the

personal account of the owner.

Persons employed in government and parastatal bodies, cooperatives, incorporated enterprises and quasi-corporate enterprises are excluded.

Also excluded are persons employed in agricultural activities, for practical reasons, and those employed in households as domestic workers. (Internation Labour Organisation, n.d.)

In LFS, Persons employed in informal sector in Rwanda are identified as all persons 16 years of age and above who were engaged in unregistered private business enterprises that did not keep written record of accounts. Unregistration meant not registered with Rwanda Revenue Authority or not paying PAYE/TPR. Domestic workers engaged by households were excluded from the classification of employment in the informal sector.

Over the past two decades, employment in the informal sector has risen rapidly in all regions in the world. Until the recent Asian economic crisis, it was only the oncerapidly-growing economies of East and Southeast Asia that experienced substantial growth of modern sector employment. However, in the wake of that crisis, most of these countries have experienced a decline in formal wage employment and a concomitant rise in informal employment. Even before the crisis, official statistics indicated that the informal sector accounted for over half of total non-agricultural employment in Latin America and the Caribbean, nearly half in East Asia, and as much as 80 percent in other parts of Asia and in Africa. (Wamuthenya, 2010)

The contribution of the informal sector not only its size is quite large. The contribution of informal sector income to total household income is significant in many regions: for example, in several African countries, informal sector income accounts for nearly 30 percent of total income and over 40 percent of total urban income. The contribution of the informal sector to gross domestic product (GDP) is probably also significant. For those countries where estimates exist, the

share of the informal sector in non-agricultural GDP is between 45 to 60 percent.

The Informal sector in Rwanda since 2000 has been undergoing a gradual transition towards a growing economy. The informal sector plays an important role in the African economies accounting for a large share of gross domestic product (GDP) and even greater share of employment. A large, unexpected growth of the informal sector would jeopardize the economic stability of developing countries.

Among the most important challenges facing governments in developing countries, including Rwanda, is the task of identifying development strategies that can generate new employment and income opportunities, and reduce under-employment and unemployment. (RUKUNDO, 2015)

After reading different publications on Labour force reports, we have seen that employment in informal sector is really significant all over the world and more particularly in Rwanda, where it stands at 71%, the results gotten from RLFS(Rwanda Labour Force Survey) February 2017 report. (NISR, 2017) Though the employment in informal sector is larger it can be increased since getting employment in informal sector is much easier than in formal sector and this will eventually decrease the unemployment rate which is higher in many countries but most importantly in Rwanda, where it stands at 16.7%(NISR, 2017), it is in this regard that within this study we will look at different factors influencing employment in informal sector and measure to which extent they affect it.

2. Problem Statement

Among the most important challenges facing governments in developing countries including Rwanda, is the task of identifying development strategies that can generate new employment and income opportunities, and reduce under-employment and unemployment.

When we talk about unemployment what clicks in people's mind is that it has increased and eventually it is correct since each year, the unemployment rate goes higher, and one can project it to be even much higher in years to come. Different policies had been tried out but they were not successful as far as getting effective and efficient solutions of this specific matter is concerned.

Knowing that informal sector employment represents roughly three quarter of the total employment in many developing countries, Rwanda included and bearing in mind that getting a job in informal sector is much easier than in formal sector, the researcher called out this specific study with the purpose of modeling different factors (educational attainment, sex, residency areas (rural and urban), age, marital status,) affecting employment in informal sector to see to which extent those factors affects employment in informal sector.

3. Justification of the Study

After reading many Labour force publications, the researcher

has observed that in many developing countries, employment in informal sector participation is higher than employment in formal sector participation and this reflects the idea that in developing countries it is easier to get an employment in informal sector than in formal sector.

Though the employment in informal sector is relatively higher than employment in formal sector, we still have the unemployment rate which is higher in developing countries. To reduce the unemployment rate, one approach that the researcher advances is to look at the main factors influencing the employment in informal sector and see what needs to be done to get a lot of people being employed in informal sector and this will definitely reduce the unemployment rate which is one of the biggest issues to be addressed.

4. Objectives of the Research

4.1 General objective

The main objective of this study will be to measure and analyze the factors influencing employment in informal sector participation.

4.2 Specific Objectives.

- 1) To determine the influence of various factors on employment in informal sector participation
- 2) To model the influence of the factors (educational attainment, sex, residency areas (rural and urban), age, marital status, poverty categories) on the employment in informal sector participation.
- 3) To test the goodness of fit of logistic regression model

5. Research Methodology

Within this part of this specific research, we will clarify the research techniques that were used to meet the objectives. This study will use secondary data got from Rwanda Labour Force survey august 2017(RLFS_August 2017).

5.1. Sample design

The sample design of the LFS is a two-stage stratified design according to which at the first stage of sampling, a stratified sample of enumeration areas from the latest population census is drawn with probabilities proportional to size measured in terms of the census number of households or census number of household members, and at the second stage of sampling, a fixed number of sample of households is selected with equal probability within each sample enumeration areas. Finally, all household members in the sample households are selected for survey interviewing. (NISR, 2017)

The scope of the survey is all persons living in private households. It excludes the institutional population permanently residing in houses such as hostels; health resorts; correctional establishments etc., as well as persons living in seasonal dwellings not covered in the survey. It also excludes workers living at their work-sites. A household is a

group of persons who live together and make common provision for food and other essentials for living. The people in the group may be related or unrelated or a combination of both. A household may consist of only one person or several persons.

5.2 Sample size

Sample size determination in most household-based surveys with multi-stage stratified design is based on the principle of first calculating the required sample size for a single «domain» assuming a simple random sample design and no non-response. A domain is a well-defined population group for which estimates with pre-determined accuracy are sought. The results are then extended to allow for non-response and deviation from simple random sampling. (NISR, 2017)

Because the sample size should be a multiple of 2x4x16 for two survey rounds per year, four rotation groups per round and 16 sample households per Primary Sampling Unit (PSU), the required annual sample size has been rounded to 18688 households. For the August 2017 round, the sample size was therefore set at Sample size of LFS August 2017 = 18688/2 = 9344 households.

5.3 Selection of Sample Enumeration Areas

The next step in sample design was the selection of the 584 sample enumeration areas. The sample was drawn in each district by probability proportional to size (pps) from the sampling frame. In practice, the sample of enumeration areas was selected from the start for an entire sequence of 18 survey rounds from August 2016 to August 2023 when in principle a new sampling frame constructed on the basis of the next population and housing census will become available. The selected sample was then assigned randomly to 18 rotation groups. Sample enumeration areas in rotation groups R1 to R4 formed the sample for LFS August 2016. The sample of enumeration areas assigned to rotation groups R3 to R6 were used for LFS August 2017. The sample enumeration areas assigned to R4 to R7 will be used for LFS February 2018, and so on. The same process will continue on each round with the addition of one rotation group per round.

5.4 Selection sample households

The sample enumeration areas were freshly listed prior to selection of the final sample of households. According to the sample design, 16 sample households were selected from the list of households in each sample enumeration area by systematic sampling with equal probability. If the list contained 16 households or less all households in the sample enumeration area were drawn in the sample. If the list contained more than 16 households, a sample of fixed size (16 households) was drawn from the list by systematic random sampling.

The purpose of listing was to ensure that every household currently residing in the sample locality has a non-zero probability of selection. Listing permits to update the sampling frame and account population movements and new household formations that have occurred since the last

preparation of the sampling frame. It aims at covering the newly constructed buildings with living quarters and taking into account demolished or vacant buildings, or transformed dwellings no longer used as living quarters, such as dwelling addresses turned to stores or workshops, or living quarters used as secondary housing units or for holidays. (NISR, 2017)

5.5 Data collection

data used within this study have been collected countrywide and from August to October 2017 in a survey called Rwanda Labour Force Survey, the specific data are the most current data of RLFS that are available on NISR website

6. Data Analysis

This study seeks to model the factors influencing the employment in informal sector, to be able to perform the required task, we will use the statistical package called STATA, and the analysis will mainly be based on different indicators that will have been brought by STATA. We guarantee that each and every variable that we will need within this study have been measured in August 2017 RLFS dataset.

6.1 Model

Given the nature of variables in this study a logistic regression model used to statistically measure the level of relationship between dependent variables and independents variables of Labour force participation in Rwanda. Logistic regression measures the relationship between the categorical dependent variable and one or more independent variables by estimating probabilities using a logistic function, which is the cumulative logistic distribution. Logistic regression can be seen as a special case of the generalized linear model and thus analogous to linear regression. The model of logistic regression, however, is based on quite different assumptions (about the relationship between dependent and independent variables) from those of linear regression. In particular, the key differences of these two models can be seen in the following two features of logistic regression. (Hosmer, 1989).

The logistic regression model is formulated as follows:

$$\log it(\pi_i) = \log\left(\frac{\pi_i}{1 - \pi_i}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$$

$i = 1, 2, 3, \dots, n$

Logistic regression can be binomial, ordinal or multinomial. Binomial or binary logistic regression deals with situations in which the observed outcome for a dependent can have only two possible types (dead vs. alive or win vs. loss).

Within this study, we will use a binary logistic regression and we will have the dependent variable either being participated in employment in informal sector or not participated in employment in informal sector.

Binary logistic model uses a dependent variable coded as 1 and 0, in this study our dependent variable will be coded as 1 if one participated in employment in informal sector and 0 if one did not participate in employment in informal sector. Logistic regression is used to predict the odds of being a case based on the values of the independent variables (predictors). The odds are defined as the probability that a particular outcome is a case divided by the probability that it is a not case. (Cox DR, 1989)

In our research, the odds will be defined as the probability that one participated employment in informal sector over the probability that one did not participate in employment in informal sector.

Like other forms of regression analysis, logistic regression makes use of one or more predictor variables that may be either continuous or categorical. Unlike ordinary linear regression, however, logistic regression is used for predicting binary dependent variables (treating the dependent variable as the outcome of a Bernoulli trial) rather than a continuous outcome.

As a predictive analysis, logistic regression model describes data and explains the relationship between one dependent variable and two or more independent variables.

In this analysis, the dependent variable is employment in informal sector participation in Rwanda and the independent variables will be:

- 1) Age category
- 2) Sex of the population (male, female)
- 3) level of education attained (none, primary, secondary and University)
- 4) Area of residency (urban, rural)
- 5) Marital status (single, married, separated)

6.2 Estimation of parameters or variables

In the model mentioned above, parameters will be interpreted as follows:

- 1) When relationship exists between two variables the correlation coefficient is positive or negative.
- 2) When there is variation in one variable related to variation to the other variable we call it the coefficient of determination.

6.3 Interpretation of logistic regression coefficients

The interpretation of the estimated regression coefficients is not as easy as in multiple regression. In logistic regression, not only is the relationship between X and Y nonlinear, but also, if the dependent variable has more than two unique values, there are several regression equations. Consider the usual case of a binary dependent variable, Y, and a single independent variable, X. Assume that Y is coded so it takes on the values 0 and 1. In this case, the logistic regression equation is

$$\ln\left(\frac{p}{1-p}\right) = b_0 + b_1 x$$

Now consider impact of a unit increase in X. The logistic regression equation becomes:

$$\ln\left(\frac{p'}{1-p'}\right) = b_0 + b_1(x+1) = b_0 + b_1 x + b_1$$

We can isolate the slope by taking the difference between these 2 equations we have:

$$b_1 = b_0 + b_1(x+1) - (b_0 + b_1 x)$$

$$b_1 = \ln\left(\frac{p'}{1-p'}\right) - \ln\left(\frac{p}{1-p}\right)$$

$$b_1 = \ln\left(\frac{\frac{p'}{1-p'}}{\frac{p}{1-p}}\right)$$

$$b_1 = \ln\left(\frac{\text{odds}'}{\text{odds}}\right)$$

That is, b_1 is the log of the ratio of the odds at X+1 and X. Removing the logarithm by exponentiating both sides gives

$$e^{b_1} = \frac{\text{odds}'}{\text{odds}}$$

The regression coefficient b_1 is interpreted as the log of the odds ratio comparing the odds after a one unit increase in X to the original odds. Note that, unlike multiple regression, the interpretation of b_1 depends on the particular value of X since the probability values, the p's, will vary for different.

To completely interpret b_1 , we must take the logarithm of the odds ratio. It is difficult to think in terms of logarithms. However, we can remember that the log of one is zero. So a positive value of b_1 indicates that the odds of the numerator are larger, while a negative value indicates that the odds of the denominator are larger. In practice, the interpretation is

based on e^{b_1} (Wu, 2016). In practice, the interpretation is

$$\text{based on } e^{b_1}$$

6.4 Hosmer-Lemeshow test

The Hosmer-Lemeshow test is a statistical test for goodness of fit for the logistic regression model. The Hosmer-Lemeshow test uses a Chi-squared statistic which is calculated as follows:

$$\chi^2_{HL} = \sum_{g=1}^G \frac{(O_g - E_g)^2}{E_g \left(1 - \frac{E_g}{n_g}\right)}$$

with O_g , E_g and n_g the observed events, expected events

and number of observations for the g^{th} risk decile group,

and G the number of groups. The test statistic follows a Chi-squared distribution with $G-2$ degrees of freedom.

A large value of Chi-squared (with small p-value < 0.05) indicates poor fit and small Chi-squared values (with larger p-value closer to 1) indicate a good logistic regression model fit. (Agresti, 2003).

6.5 Effects measurement of all different factors on employment in informal sector participation using logistic regression model

Table 1: Employment in informal sector participation model

Iteration 0: log likelihood = -3955.811					
Iteration 1: log likelihood = -2982.0965					
Iteration 2: log likelihood = -2951.9246					
Iteration 3: log likelihood = -2951.7779					
Iteration 4: log likelihood = -2951.7779					
Logistic regression		Number of obs =	6315		
Log likelihood = -2951.7779		LR chi2(5) =	2008.07		
		Prob > chi2 =	0.0000		
		Pseudo R2 =	0.2538		
IS2_LOGISTIC	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
A01	.2077963	.0685467	3.03	0.002	.0734472 .3421453
Code_UR	.3757036	.0701522	5.36	0.000	.2382078 .5131995
marital_status	-.0656324	.0393257	-1.67	0.095	-.1427094 .0114446
age3	-.1861422	.0435304	-4.28	0.000	-.2714603 -.1008241
attained	-.8758241	.0252851	-34.64	0.000	-.9253821 -.8262661
_cons	2.837221	.2208038	12.85	0.000	2.404454 3.269989

6.6 Interpretation of the logistic regression model variables

Researcher considers the STATA output of logistic regression model, He analyses and interprets the effect of each fact taken as independent variable on employment in informal sector participation.

Considering the STATA output above given, the logistic regression model is given as follows:

$$y_i = \log \text{it}(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right) = 2.838 + 0.208x_1 + 0.376x_2$$

$$-0.066x_3 - 1.861x_4 - 0.876x_5$$

Where $i:1,2,3,4,5$ and y_i is the dependent variable for i factor while x_1, x_2, x_3, x_4 and x_5 are independent factors.

Analyzing the above model, the researcher observed that factor sex (A01 Variable in STATA output) has a statistical significant impact on employment in informal sector participation. The female persons are most likely to join the employment in informal sector than men and the same goes for the residency area factor, people living in rural areas are most likely to be part of the employment in informal sector than people in urban areas.

On the hand the factor age category has a negative parameter (-1.861) this means that as we move from the first category (16-30 years old) to the last category (65+ years old) of age

interval people are less likely to join the employment in informal sector.

Level of education attained also has a statistical significant impact on the employment in informal sector, its negative parameter (-0.876) means that as people increases their levels of education they are less likely to join the employment in informal sector.

Only the marital status factor has shown no statistical impact on the employment in informal sector due to its p-value $0.095 > 0.05$.

7. Summary, Conclusion and Recommendation

7.1 Summary

This research that we conducted all of the country reveals that among different factors that we though could have effects on employment in informal sector participation, four of those have shown statistical impact on employment in informal sector participation and those are sex of the population, people's level of education, age category of the population and residency area of the population (urban or rural).

We have observed that women are more likely to be part of the employment in informal sector than men. The researcher also clarifies that people are less likely to join the employment in informal sector as they increase their studies. it has been proven that people living in rural areas are more likely to join the employment in informal sector than people in urban areas of Rwanda. It is also observed that Persons are less likely to go for employment in informal sector as they grow up in age.

7.2 Conclusion

This research has been analyzing the effects of different factors on employment in informal sector participation using Labor force survey data collected from august to October 2017. In the analysis, the researcher analyzed the effects of each factor on the employment in informal sector participation and it comes out that among all factors considered only four of them have shown a statistically significant impact on employment in informal sector and the remaining showed an effect on it but which are not statistically significant.

After analyzing the logistic regression model coefficients, it is evident that women are more likely to join the employment in informal sector than men, we relate this finding to the fact that a lot of men neglect small jobs and small businesses yet some of them do not have any other source of income.

The analysis part of his research also showed that people living in rural areas are more likely to be involved in small businesses and small jobs considered as employment in informal sector compared to persons living in urban areas of Rwanda. We may attribute this finding to the reality on the ground that making a small business in urban areas of

Rwanda especially in Kigali is difficult and requires a lot of money. A good example is the street vendors who are being eliminated from streets, they are pushed to make their small businesses in the houses which are expensive yet their capital cannot strength that far.

It has been observed also that people with lower levels of educations are more likely to join the employment in informal sector. This is can also be justified by the mentality of many intellectuals that people in high levels of education should occupy bureaucracy positions forgetting that those jobs are very limited in number.

The researcher observes also that people are less likely to join the employment in informal sector as we move from first category of (16-30 years old) to the last category (65+ years old). This can be understood by the fact that when one is getting older he fears the employment in informal sector since most of the time are tiresome and people involved in employment in informal sector are informally employed.

7.3 Recommendations

Many countries all over the world took different economic policies to address the problem of unemployment which increases very rapidly. In line with this context, Rwanda as a country has been trying out different policies to reduce the unemployment rate which stands at 17.8%. (NISR, Labour Force Survey August 2017, 2017)

Rwanda has very recently initiated the NST program within this specific program, to reduce the unemployment, there will be a creation of decent jobs for economic development and poverty reduction.

Knowing that jobs in formal sector are very limited in number and from the fact that getting a job in formal sector depend upon many factors which disqualify many people. Governments of Rwanda should put more efforts through its employment policies to encourage and motivate people by all means to look for employment in informal sector and this will definitely reduce the unemployment in Rwanda. However, thought getting the employment in informal sector is much easier than in formal sector as discussed previously many factors influence the employment in informal sector. This research measured impact of different factors in the employment in informal sector and the results have been shown in the result part of this research.

The researcher reveals that people living in rural areas are more likely to join the employment in informal sector than people in urban areas of Rwanda, in line with this finding the research suggests that government of Rwanda should facilitate the employment in informal sector by allowing street vending, under the table small businesses. in urban areas of Rwanda especially in Kigali city where a lot of people in informal sector employment are found.

The researcher suggests also the government of Rwanda to encourage or mobilize persons with high level of education; those persons with secondary education or even universities to be part of the employment in informal sector instead of

spending many years looking for employment in formal sectors, or neglecting the employment in informal sector since it is also a source of income. The researcher also recommends the government of Rwanda to mobilize the men to join the employment in informal sector since the research reveals that men are less likely to join the employment in informal sector than female and this is because a lot of people(men) neglect the informal sector employment.

Reaction on the above recommendations by government of Rwanda through its employment policies and programs will result in having more people in employment in informal sector which eventually reduces the unemployment rate in Rwanda.

References

- [1] Agresti, A. (2003). Logit models for multinomial responses. *Categorical Data Analysis, Second Edition*, 267--313.
- [2] Cox DR, S. E. (1989). *D.R.Cox, E.J. Snell*. Hosmer, D. W. (1989). *Applied logistic regression* . New york : Wiley Library .
- [3] *Internation Labour Organisation*. (n.d.). Retrieved from <http://www.ilo.org>.
- [4] Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European sociological review*, 67--82.
- [5] NISR. (2017). *Labour Force Survey August 2017*. Kigali : NISR.
- [6] NISR. (2017). *RWANDA LABOUR FORCE SURVEY. KIGALI: NISR*.
- [7] RUKUNDO, J. B. (2015). Understanding Informal sector Employment in Rwanda. *Online International Research Journal*.
- [8] Wamuthenya, W. R. (2010). Determinants of Employment in the formal and informal sectors of the urban areas of kenya . *AERC Research Paper 194*.
- [9] Wu, Q. a. (2016). Exploratory multinomial logit model--based driver injury severity analyses for teenage and adult drivers in intersection-related crashes. *Traffic injury prevention*, 413--422.